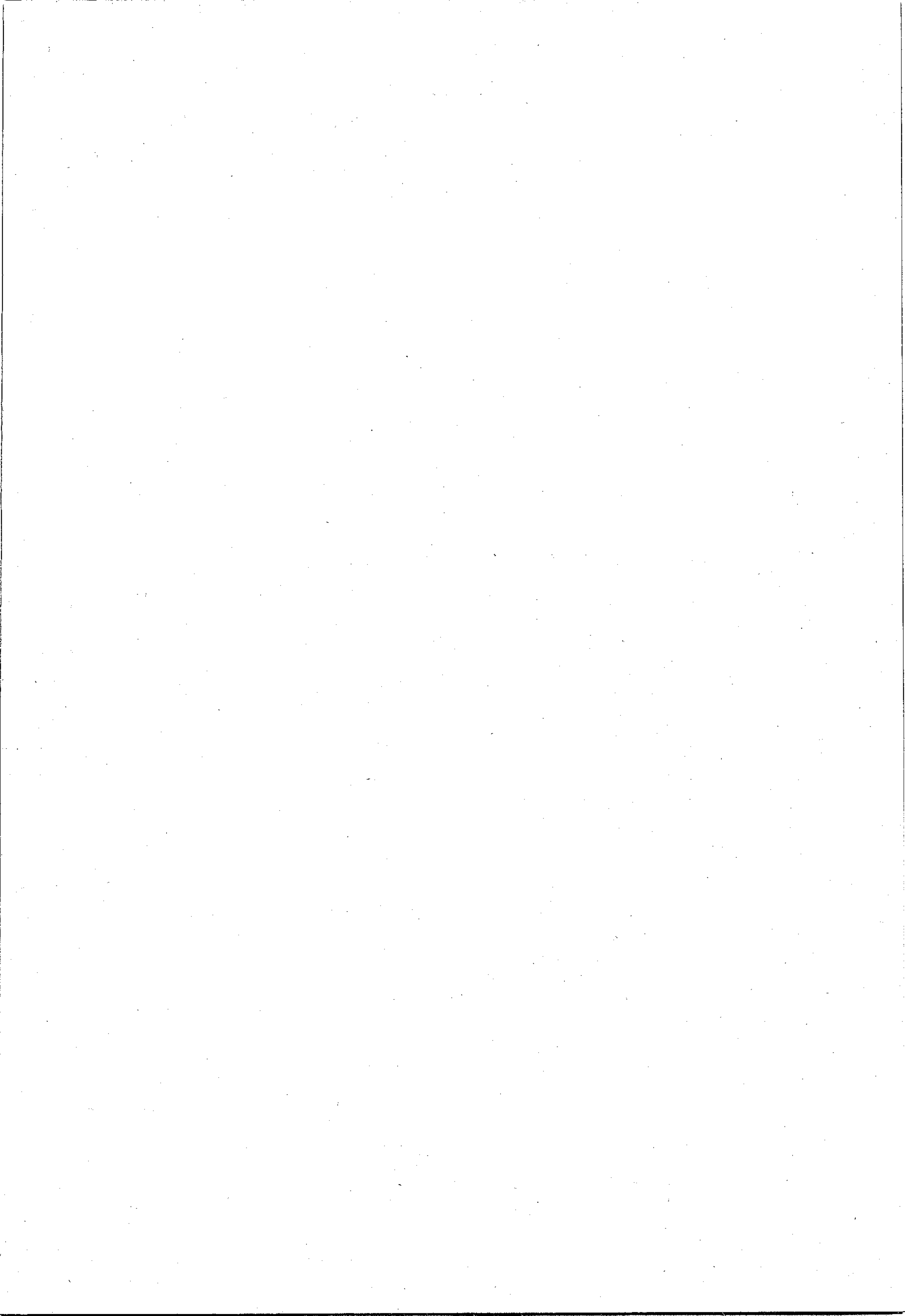
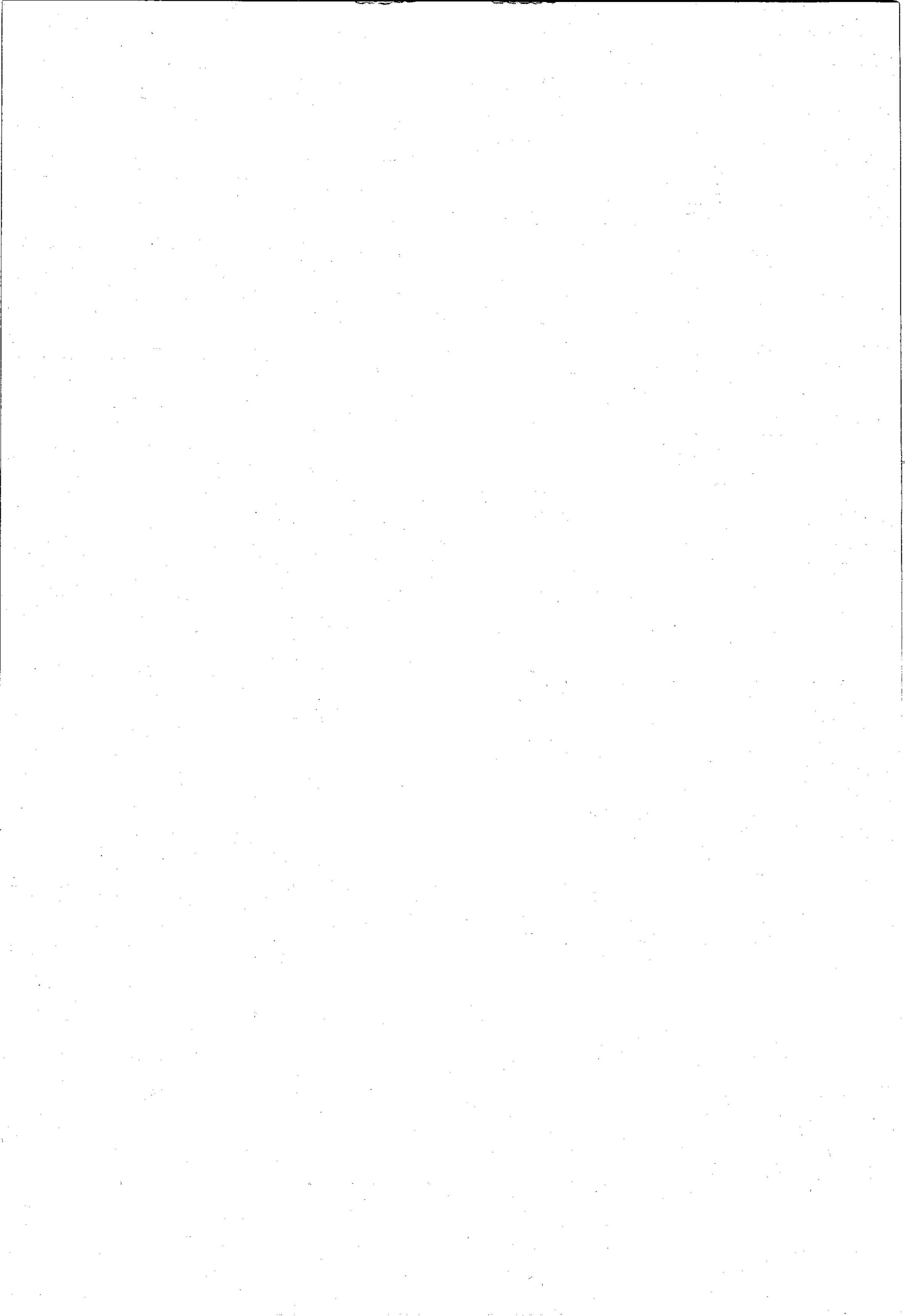
A. N. Kolmogorov S. V. Fomin

Reelle Funktionen und Funktionalanalysis

A. N. Kolmogorov
S. V. Fomin
Reelle Funktionen und Funktionalanalysis





A. N. Kolmogorov / S. V. Fomin
Reelle Funktionen und Funktionalanalysis

Hochschulbücher für Mathematik

Herausgegeben von H. Grell, K. Maruhn und W. Rinow
Band 78

Reelle Funktionen und Funktionalanalysis

von A. N. Kolmogorov und S. V. Fomin

Mit 24 Abbildungen



VEB Deutscher Verlag der Wissenschaften
Berlin 1975

А. Н. Колмогоров
С. В. Фомин

Элементы теории функций и функционального анализа
Издание третье, переработанное
„Наука", Москва 1972

Übersetzung aus dem Russischen: Dr. D. Freitag, Dipl.-Math. H. Palme, Dr. B. Stöckert
Wissenschaftliche Redaktion: Prof. Dr. H. Triebel

Verlagslektoren: Dipl.-Math. Brigitte Mai, Dipl.-Math. Gesine Reiher
Verlagshersteller: Barbara Harnack
Gestalter für Einband und Schutzumschlag: Rudolf Wendt

Printed in the German Democratic Republic
Lizenz-Nr. 206 · 435/104/75
Gesamtherstellung: VEB Druckhaus „Maxim Gorki", 74 Altenburg
LSV 1034
Bestellnummer 570 227 6
EVP 55,— Mark

Aus dem Vorwort zur zweiten russischen Auflage

Die erste Auflage dieses Buches erschien in den Jahren 1954 und 1960 in zwei einzelnen Bänden, deren Inhalt dem Programm der Vorlesung „Analysis III" entsprach, wie sie Ende der vierziger Jahre an der Mechanisch-mathematischen Fakultät der Moskauer Lomonossov-Universität gehalten wurde. In diesem Programm waren Elemente der Maßtheorie, der Theorie der reellen Funktionen, der Integralgleichungen, der Funktionalanalysis und später auch der Variationsrechnung enthalten. In der folgenden Zeit wurde diese Vorlesung, die an der Moskauer Lomonossov-Universität zuerst von A. N. Kolmogorov und danach von anderen Dozenten, darunter S. V. Fomin, gehalten wurde, auch von anderen Universitäten übernommen.

Eine Zusammenfassung der einzelnen Vorlesungen über die Theorie der reellen Funktionen, über Integralgleichungen und über Variationsrechnung zu einer einzigen entfachte seinerzeit große Diskussionen. Die neue Vorlesung hatte die Aufgabe, einerseits die innere Logik der Entwicklung der Mengenlehre, der allgemeinen Theorie der stetigen Abbildungen in metrischen und topologischen Räumen, der linearen Räume, der Funktionale und Operatoren auf diesen Räumen, der reinen Maß- und Integrationstheorie auf allgemeinem „Mengen mit Maß" darzustellen; andererseits hatte sie die Aufgabe, bei der Behandlung dieser abstrakten Gebiete der Mathematik nicht die Probleme der klassischen und der angewandten Analysis aus den Augen zu verlieren.

Bei der Lösung dieser Aufgabe gaben wir im Buch dem abstrakten Aufbau den Vorrang. Von der allgemeinen Mengenlehre (Kapitel 1) kann man zu den metrischen und topologischen Räumen und ihren stetigen Abbildungen (Kapitel 2) oder unmittelbar zu den Mengen mit Maß und der Integration auf ihnen (Kapitel 5) übergehen. In den Kapiteln 3 und 4 werden lineare Räume sowie lineare Funktionale und Operatoren auf ihnen untersucht. Von diesen Kapiteln ist auch ein direkter Übergang zu Kapitel 10 (nichtlineare differenzierbare Operatoren und Funktionale) möglich. Im Kapitel 7 werden lineare Räume integrierbarer Funktionen untersucht. Nur in den Kapiteln 6 und 8 bilden die Funktionen einer reellen Veränderlichen den eigentlichen Gegenstand der Betrachtung.

Obwohl in unserem Buch in erster Linie allgemeine Begriffe der Theorie der reellen Funktionen und der Funktionalanalysis dargestellt werden, wird der Leser in fast allen Kapiteln auf die angrenzenden klassischen Probleme aufmerksam gemacht. Die Aufnahme der Kapitel 6 (Theorie der Differentiation), 8 (trigonometrische

Reihen und Fouriersches Integral) und 9 (lineare Integralgleichungen) führte dazu, daß unser Buch jetzt das gesamte an der Moskauer Lomonossov-Universität gebotene Programm „Analysis III" mit Ausnahme der Variationsrechnung umfaßt. Wir haben dieses Gebiet in das Buch nicht aufgenommen, sondern beschränken uns im Kapitel 10 auf einführende Darlegungen über nichtlineare Funktionalanalysis.

In der neuen Auflage nimmt — wie auch in der ersten — die allgemeine Maßtheorie einen bedeutenden Platz ein. Obwohl in der letzten Zeit ziemlich viele Darstellungen erschienen, in denen der Integrationstheorie das Daniellsche Schema, das nicht den Apparat der Maßtheorie benutzt, zugrunde gelegt ist, glauben wir, daß die Maßtheorie unabhängig von der Einführung des Integralbegriffs für sich sehr wichtig ist und Gegenstand einer Vorlesung sein sollte.

Die Aufnahme neuer Kapitel hat den Umfang des Buches stark vergrößert. Alte Kapitel wurden wesentlich überarbeitet und durch neue Abschnitte ergänzt (z. B. über Ordnungstypen, transfinite Zahlen und topologische Räume, Distributionen u. a.).

Bei der Bearbeitung unseres Buches und der Aufnahme neuer Teile waren wir bemüht, den — wie wir glauben — verhältnismäßig elementaren Stil der Darlegung der ersten Auflage beizubehalten. Wir hoffen, daß das Buch neben anderen Werken seinen natürlichen Platz in der Hochschulbuchliteratur finden wird. Wir weisen in diesem Zusammenhang insbesondere auf das Buch von G. E. Šilov „Mathematische Analysis, Spezieller Kurs" hin, in dem das Interesse mehr der analytischen Seite der Gegenstände und nicht so sehr den metrischen und topologischen Räumen, Maßen usw. als selbständigen Objekten gilt.

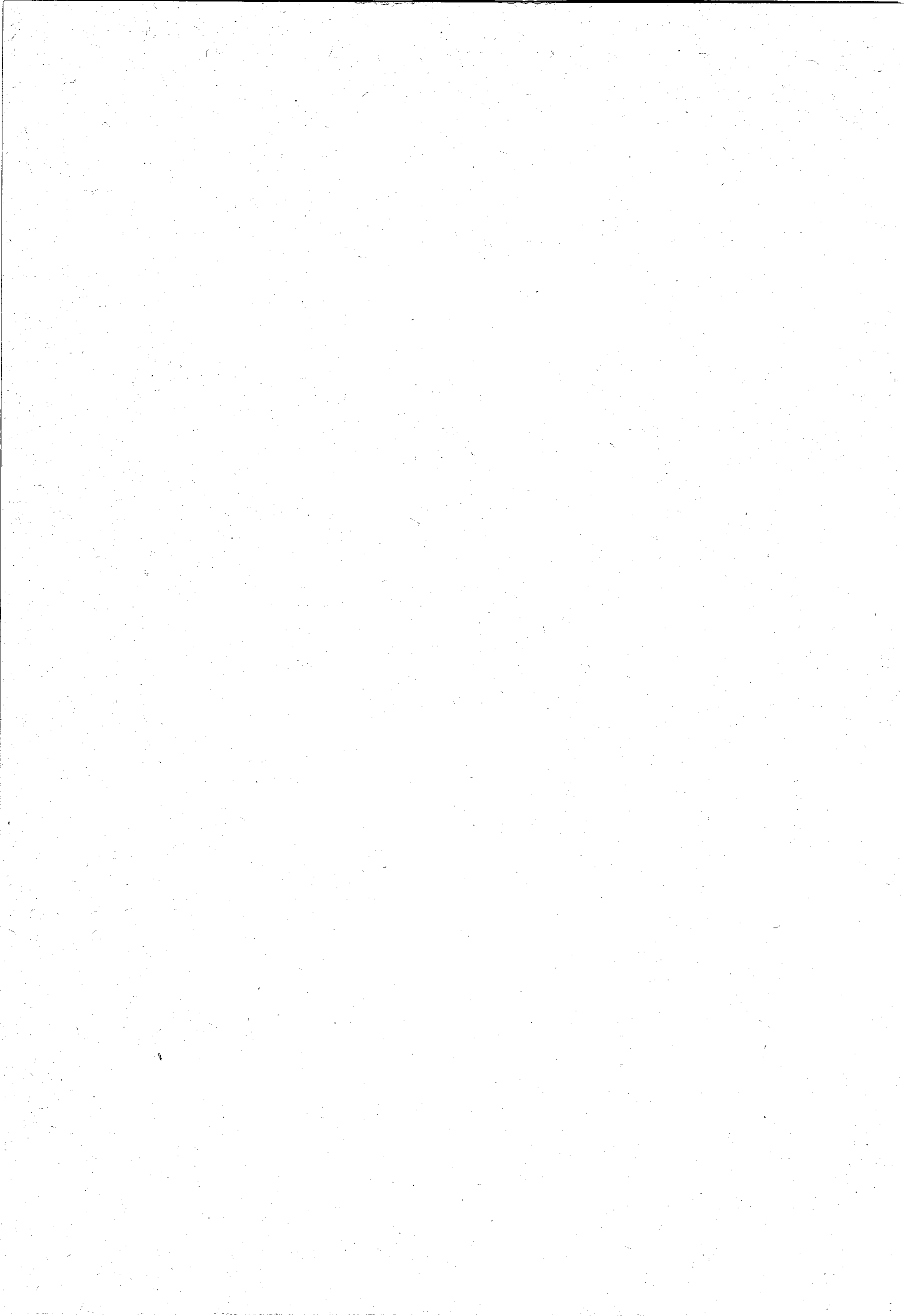
A. Kolmogorov
S. Fomin

Vorwort zur dritten russischen Auflage

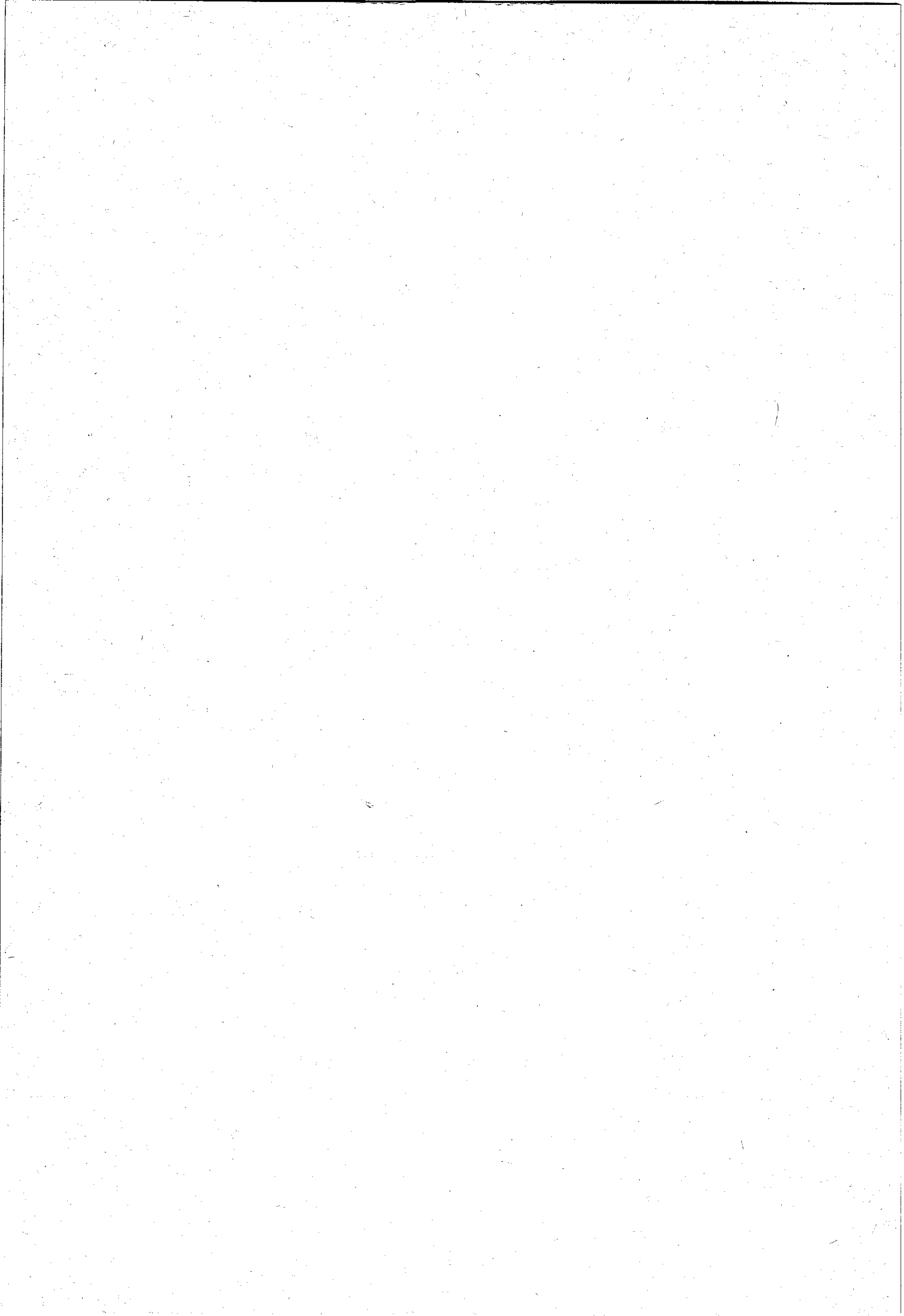
Bei der Vorbereitung der neuen Auflage haben wir die Grundkonzeption des Buches beibehalten und waren bestrebt, seinen Umfang nicht zu vergrößern. Der gesamte Text wurde erneut durchgesehen und überarbeitet, wobei wir wesentlich von F. V. Širokov unterstützt wurden. In den Kapiteln 1 und 4 haben wir einige Umstellungen und Veränderungen vorgenommen, die nach unserer Meinung den Übergang von den einfacheren Begriffen zu den komplizierteren (z. B. von den Banachräumen zu den allgemeinen Räumen im Kapitel 4) erleichtern. Wesentlich überarbeitet wurde die Darstellung der Maßtheorie (Kapitel 5).

Seit einigen Jahren werden in die Vorlesung „Analysis III" oft Elemente der Theorie der Banachschen Algebren und der Spektralanalysis aufgenommen, so daß wir es für zweckmäßig hielten, einen von V. M. Tichomirov geschriebenen Anhang über diese Fragen in unser Buch aufzunehmen.

A. Kolmogorov
S. Fomin



Inhalt



1. Elemente der Mengenlehre

1.1. Der Begriff der Menge. Operationen mit Mengen

1.1.1. Grundlegende Definitionen. In der Mathematik stößt man auf *Mengen* mannigfaltigster Art. Man kann von der Menge der Kanten eines Polyeders, der Menge der Punkte auf einer Geraden, der Menge der natürlichen Zahlen usw. sprechen. Der Begriff der Menge ist derart allgemein, daß es schwierig ist, ihm irgendeinen Sinn zu geben, der nicht einfach auf eine Vertauschung des Wortes „Menge" mit seinen Synonymen Gesamtheit, Sammlung von Elementen usw. hinausläuft.

Die Rolle des Mengenbegriffs in der modernen Mathematik erklärt sich nicht nur daraus, daß die Mengenlehre als solche in der jetzigen Zeit zu einer umfang- und inhaltsreichen Disziplin geworden ist. In der Hauptsache resultiert sie aus dem Einfluß, den die Mengenlehre seit ihrem Entstehen am Ende des vergangenen Jahrhunderts auf die gesamte Mathematik ausgeübt hat und noch ausübt. Ohne eine vollständige Darstellung aller möglichen Probleme dieser Theorie zu geben, wollen wir hier lediglich die grundlegenden Bezeichnungen einführen und mit den grundlegenden mengentheoretischen Begriffen vertraut machen, die wir später verwenden werden.

Mengen werden wir mit großen Buchstaben $A, B, \ldots$ und ihre Elemente mit kleinen Buchstaben $a, b, \ldots$ bezeichnen. Die Aussage „Das Element a gehört zu der Menge A" wird symbolisch durch $a \in A$ oder $A \ni a$ dargestellt; die Schreibweise $a \notin A$ (oder $A \not\ni a$) bedeutet, daß das Element a nicht zu A gehört. Wenn alle Elemente der Menge A auch einer Menge B angehören(wobei der Fall $A = B$ nicht ausgeschlossen ist), so nennen wir A *Teilmenge* oder *Untermenge* von B und schreiben $A \subset B$. Beispielsweise bilden die ganzen Zahlen eine Teilmenge der Menge aller reellen Zahlen.

Manchmal wissen wir nicht von vornherein, ob eine Menge (zum Beispiel die Menge der Wurzeln einer gegebenen Gleichung) auch nur ein Element enthält. Deshalb ist es zweckmäßig, den Begriff der *leeren* Menge einzuführen, d. h. die Menge, die kein einziges Element enthält. Wir werden sie mit dem Symbol $\emptyset$ bezeichnen. Jede Menge enthält $\emptyset$ als Teilmenge.

1.1.2. Operationen mit Mengen. Es seien A und B beliebige Mengen. Als ihre *Summe* oder *Vereinigung* $C = A \cup B$ bezeichnet man diejenige Menge, die aus allen Elementen besteht, die mindestens einer der beiden Mengen A und B angehören (Abb. 1).

Analog definiert man die Vereinigung von beliebig (endlich oder unendlich) vielen Mengen: Wenn A_α beliebige Mengen sind, so ist ihre Vereinigung $\bigcup_\alpha A_\alpha$ die Gesamtheit aller Elemente, von denen jedes zu mindestens einer der Mengen A_α gehört.

Als *Durchschnitt* $C = A \cap B$ der beiden Mengen A und B bezeichnen wir diejenige Menge, die aus allen Elementen besteht, die sowohl zu A als auch zu B gehören (Abb. 2). Beispielsweise besteht der Durchschnitt der Menge aller geraden Zahlen

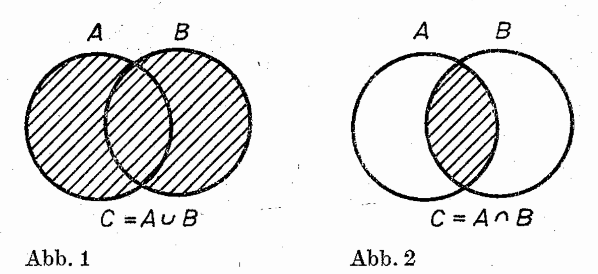


Abb. 1 Abb. 2

und der Menge aller Zahlen, die durch 3 teilbar sind, aus allen ganzen Zahlen, die ohne Rest durch 6 teilbar sind. Als Durchschnitt beliebig (endlich oder unendlich) vieler Mengen A_α bezeichnen wir die Gesamtheit $\bigcap_\alpha A_\alpha$ derjenigen Elemente, die jeder der Mengen A angehören.

Die Operationen der Vereinigung und der Durchschnittsbildung von Mengen sind nach Definition kommutativ und assoziativ, d. h.

$$A \cup B = B \cup A, \quad (A \cup B) \cup C = A \cup (B \cup C),$$

$$A \cap B = B \cap A, \quad (A \cap B) \cap C = A \cap (B \cap C).$$

Außerdem gilt das Distributivgesetz:

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C), \tag{1}$$

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C). \tag{2}$$

Wir wollen beispielsweise die erste dieser Gleichungen nachprüfen.[1]) Das Element x gehöre zu der Menge, die auf der linken Seite der Gleichung (1) steht, d. h. $x \in (A \cup B) \cap C$. Das bedeutet, daß x der Menge C und außerdem mindestens einer der Mengen A oder B angehört. Dann gehört aber x mindestens zu einer der Mengen $A \cap C$ oder $B \cap C$, d. h., x ist Element der rechten Seite der betrachteten Gleichung. Ist umgekehrt $x \in (A \cap C) \cup (B \cap C)$, so gilt $x \in A \cap C$ oder $x \in B \cap C$. Folglich ist $x \in C$, und außerdem gehört x zu A oder B, d. h. $x \in A \cup B$. Damit ist $x \in (A \cup B) \cap C$, und (1) ist bewiesen. Analog prüft man (2) nach.

---

[1]) Die Gleichheit zweier Mengen $A = B$ ist als identische Gleichheit zu verstehen, d. h, daß jedes Element der Menge A auch zu B gehört und umgekehrt. Mit anderen Worten, die Gleichung $A = B$ ist gleichbedeutend damit, daß $A \subset B$ und $B \subset A$ erfüllt sind.

Wir definieren nun den Begriff der Differenz von Mengen. Man bezeichnet als *Differenz* $C = A \setminus B$ der Mengen A und B die Gesamtheit derjenigen Elemente von A, die nicht zu B behören (Abb. 3). Dabei muß nicht unbedingt vorausgesetzt werden, daß $A \subset B$ gilt. An Stelle von $A \setminus B$ wird oft auch $A - B$ geschrieben.

Manchmal (zum Beispiel in der Maßtheorie) ist es sinnvoll, die sogenannte *symmetrische Differenz* zweier Mengen A und B zu betrachten, die als Vereinigung der

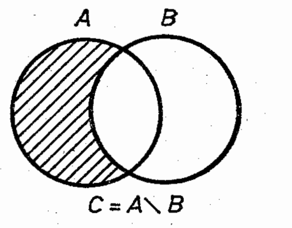


Abb. 3

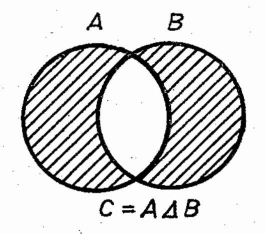


Abb. 4

Differenzen $A \setminus B$ und $B \setminus A$ definiert ist (Abb. 4). Die symmetrische Differenz C der Mengen A und B wollen wir mit $A \triangle B$ bezeichnen. Nach Definition ist somit

$$A \triangle B = (A \setminus B) \cup (B \setminus A).$$

Aufgabe. Man beweise die Gleichung

$$A \triangle B = (A \cup B) \setminus (A \cap B).$$

Oft ist es notwendig, ein System von Mengen zu betrachten, die Teilmengen einer gewissen Grundmenge S sind, zum Beispiel verschiedene Punktmengen auf der Zahlengeraden. In diesem Fall bezeichnet man die Differenz $S \setminus A$ als das *Komplement* der Menge A und schreibt $\complement A$ oder A'.

In der Mengenlehre und ihren Anwendungsgebieten spielt das sogenannte *Dualitätsprinzip* eine äußerst wichtige Rolle, das auf den folgenden beiden Relationen basiert:

1. *Das Komplement der Vereinigung ist gleich dem Durchschnitt der Komplemente:*

$$S \setminus \bigcup_\alpha A_\alpha = \bigcap_\alpha (S \setminus A_\alpha). \tag{3}$$

2. *Das Komplement des Durchschnitts ist gleich der Vereinigung der Komplemente:*

$$S \setminus \bigcap_\alpha A_\alpha = \bigcup_\alpha (S \setminus A_\alpha). \tag{4}$$

Das Dualitätsprinzip besteht darin, daß man aus einem beliebigen Satz über ein System von Teilmengen einer fixierten Grundmenge S automatisch einen anderen — dualen — Satz auf dem Wege gewinnt, daß man alle betrachteten Mengen durch ihre Komplemente, Vereinigungen von Mengen durch ihre Durchschnitte und Durchschnitte durch ihre Vereinigungen ersetzt. Als Beispiel für die Anwendung dieses Prinzips kann die Herleitung des Satzes 3′ aus Satz 3 in 2.2. dienen.

Wir wollen nun den Beweis für die Beziehung (3) führen.

Es sei $x \in S \setminus \bigcup_\alpha A_\alpha$. Das bedeutet, daß x nicht zu der Vereinigung $\bigcup_\alpha A_\alpha$, also zu keiner der Mengen A_α gehört. Folglich gehört x zu jedem der Komplemente $S \setminus A_\alpha$, und damit gilt $x \in \bigcap_\alpha (S \setminus A_\alpha)$, Ist umgekehrt $x \in \bigcap_\alpha (S \setminus A_\alpha)$, d. h., daß x zu jedem $S \setminus A_\alpha$ gehört, dann liegt x in keiner der Mengen A_α, d. h., x gehört auch nicht zu ihrer Vereinigung $\bigcup_\alpha A_\alpha$, und folglich gilt $x \in S \setminus \bigcup_\alpha A_\alpha$. Gleichung (3) ist damit bewiesen. Die Beziehung (4) wird analog bewiesen (der Beweis sei dem Leser überlassen).

Die Bezeichnung „symmetrische Differenz" für die Operation $A \triangle B$ ist nicht in jeder Hinsicht glücklich gewählt. Diese Operation ist in vielem der Vereiniging $A \cup B$ analog. $A \cup B$ bedeutet nämlich, daß wir durch ein *nicht ausschließendes* „oder" zwei Aussagen miteinander verbinden: „Das Element gehört zu A" und „Das Element gehört zu B"; $A \triangle B$ bedeutet, daß diese Aussagen durch ein *ausschließendes* „oder" verbunden werden: Ein Element x gehört zu $A \triangle B$ dann und nur dann, wenn es entweder *nur* zu A oder *nur* zu B gehört. Die Menge $A \triangle B$ könnte man besser als „Vereinigung modulo 2" der Mengen A und B bezeichnen. (Es wird eine Vereinigung dieser beiden Mengen gebildet, aber die Elemente, die dadurch doppelt auftreten, werden weggelassen.)

1.2. Abbildungen. Klasseneinteilungen

1.2.1. Abbildungen von Mengen. Der allgemeine Funktionsbegriff. In der Analysis wird der Begriff der Funktion auf folgende Weise eingeführt. Es sei X irgendeine Menge auf der Zahlengeraden. Man sagt, daß auf dieser Menge *eine Funktion f definiert* ist, wenn jeder Zahl $x \in X$ eine wohldefinierte Zahl $y = f(x)$ zugeordnet ist. Dabei nennt man X den *Definitionsbereich* der gegebenen Funktion und die Gesamtheit Y aller Werte, die diese Funktion annimmt, ihren *Wertevorrat.*

Betrachtet man jedoch statt Zahlenmengen irgendwelche anderen Mengen, so kommen wir zum allgemeinen Funktionsbegriff. Es seien M und N zwei beliebige Mengen. Man sagt, daß auf M eine Funktion f definiert ist, die Werte aus N annimmt, wenn jedem Element $x \in M$ ein und nur ein Element y aus N zugeordnet wird. Im Fall von Mengen beliebiger Art (wie übrigens auch im Fall von Zahlenmengen) bedient man sich oft statt des Terminus „Funktion" auch des Terminus „Abbildung", wobei man dann von einer Abbildung der einen Menge in die andere spricht. Wenn man Mengen M und N von spezieller Art wählt, so entstehen spezielle Typen von Funktionen, die als „Vektorfunktionen", „Maße", „Funktionale", „Operatoren" usw. bezeichnet werden. Wir werden ihnen im Weiteren noch begegnen.

Zur Bezeichnung einer Funktion (Abbildung) von M in N werden wir uns oft der Schreibweise $f\colon M \to N$ bedienen.

Ist a ein Element aus M, so nennt man das ihm entsprechende Element $b = f(a)$ aus N sein (durch die Abbildung f vermitteltes) *Bild.* Die Gesamtheit aller jener Elemente a aus M, deren Bild ein gegebenes Element $b \in N$ ist, nennt man das *Urbild* (oder besser *volles Urbild*) des Elementes b und bezeichnet es mit $f^{-1}(b)$. Es sei A

eine Teilmenge von M; die Gesamtheit $\{f(a): a \in A\}$ aller Elemente der Gestalt $f(a)$ mit $a \in A$ nennt man das *Bild von* A und bezeichnet es mit $f(A)$. Seinerseits definiert man für jede Menge B aus N ihr (volles!) Urbild $f^{-1}(B)$ durch: $f^{-1}(B)$ ist die Gesamtheit aller jener Elemente aus M, deren Bilder zu B gehören. Es kann passieren, daß kein einziges Element b aus B ein Urbild hat, dann ist das Urbild $f^{-1}(B)$ die leere Menge.

Wir werden uns hier auf die Betrachtung der allgemeinsten Eigenschaften von Abbildungen beschränken.

Es wird folgende Terminologie eingeführt. Wir sagen, daß f eine Abbildung der Menge M auf die Menge N ist, wenn $f(M) = N$ gilt; eine solche Abbildung nennt man auch *surjektiv*. Im allgemeinen Fall, d. h. wenn $f(M) \subset N$ gilt, sagt man, f sei eine Abbildung von M „in" N.

Wenn für zwei beliebige verschiedene Elemente x_1 und x_2 aus M ihre Bilder $y_1 = f(x_1)$ und $y_2 = f(x_2)$ ebenfalls verschieden sind, so nennt man f eine *injektive* (oder *eineindeutige*) Abbildung.

Wir wollen jetzt grundlegende Eigenschaften von Abbildungen beweisen.

Satz 1. *Das Urbild der Vereinigung zweier Mengen ist gleich der Vereinigung ihrer Urbilder:*

$$f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B).$$

Beweis. Das Element x gehöre zur Menge $f^{-1}(A \cup B)$. Das bedeutet, daß $f(x) \in A \cup B$, d. h. $f(x) \in A$ oder $f(x) \in B$ ist. Dann gehört aber x mindestens einer der Mengen $f^{-1}(A)$ oder $f^{-1}(B)$ an, d. h. $x \in \big(f^{-1}(A) \cup f^{-1}(B)\big)$. Ist umgekehrt $x \in \big(f^{-1}(A) \cup f^{-1}(B)\big)$, so gehört x mindestens einer der Mengen $f^{-1}(A)$ oder $f^{-1}(B)$ an, d. h., $f(x)$ gehört zu mindestens einer der Mengen A oder B, folglich ist $f(x) \in A \cup B$ und damit $x \in f^{-1}(A \cup B)$.

Satz 2. *Das Urbild des Durchschnitts zweier Mengen ist gleich dem Durchschnitt ihrer Urbilder:*

$$f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B).$$

Beweis. Ist $x \in f^{-1}(A \cap B)$, so gilt $f(x) \in A \cap B$, d. h.

$$f(x) \in A \quad \text{und} \quad f(x) \in B;$$

folglich ist $x \in f^{-1}(A)$ und $x \in f^{-1}(B)$, also $x \in \big(f^{-1}(A) \cap f^{-1}(B)\big)$. Ist umgekehrt $x \in \big(f^{-1}(A) \cap f^{-1}(B)\big)$, d. h. $x \in f^{-1}(A)$ und $x \in f^{-1}(B)$, so ist $f(x) \in A$ und $f(x) \in B$. Mit anderen Worten, es ist $f(x) \in A \cap B$. Folglich gilt $x \in f^{-1}(A \cap B)$.

Die Sätze 1 und 2 bleiben auch gültig für Vereinigungen und Durchschnitte einer beliebigen (endlichen oder unendlichen) Anzahl von Mengen; dasselbe gilt auch für den folgenden Satz.

S a t z 3. *Das Bild der Vereinigung zweier Mengen ist gleich der Vereinigung ihrer Bilder:*

$$f(A \cup B) = f(A) \cup f(B).$$

B e w e i s. Ist $y \in f(A \cup B)$, so heißt das, daß $y = f(x)$ gilt, wobei x mindestens einer der Mengen A und B angehört. Folglich gehört $y = f(x)$ zu $f(A) \cup f(B)$. Ist umgekehrt $y \in \big(f(A) \cup f(B)\big)$, so ist $y = f(x)$, wobei x mindestens einer der Mengen A und B angehört, d. h., es ist $x \in A \cup B$ und folglich $y = f(x) \in f(A \cup B)$.

Wir bemerken, daß *das Bild des Durchschnitts zweier Mengen im allgemeinen nicht mit dem Durchschnitt ihrer Bilder übereinstimmt.* Stellt beispielsweise die betrachtete Abbildung eine Projektion der Ebene auf die x-Achse dar, so schneiden sich die Strecken

$$0 \leqq x \leqq 1, \qquad y = 0,$$

$$0 \leqq x \leqq 1, \qquad y = 1,$$

nicht, während ihre Bilder zusammenfallen.

A u f g a b e. Man zeige, daß das Urbild des Komplements gleich dem Komplement des Urbildes ist. Gilt die analoge Aussage auch für die Bilder der Komplemente?

1.2.2. Klasseneinteilung. Äquivalenzrelationen. Im Zusammenhang mit den verschiedensten Fragen stößt man auf Einteilungen von Mengen in paarweise disjunkte Teilmengen. Zum Beispiel kann man die Ebene (wenn man sie als Punktmenge betrachtet) in Geraden zerlegen, die parallel zur x-Achse verlaufen; den dreidimensionalen Raum kann man sich vorstellen als Gesamtheit konzentrischer Kugeloberflächen verschiedener Radien (beginnend mit $r = 0$); die Bewohner einer bestimmten Stadt kann man in Gruppen nach Geburtsjahren einteilen usw.

Jedesmal, wenn eine gewisse Menge M auf irgendeine Art als Vereinigung paarweise disjunkter Teilmengen betrachtet wird, sprechen wir von einer *Einteilung der Menge M in Klassen.*

Für gewöhnlich hat man es mit Klasseneinteilungen zu tun, die nach einem Kriterium konstruiert sind, mit dessen Hilfe die Elemente der Menge M in Klassen zusammengefaßt werden. Beispielsweise kann man die Menge aller Dreiecke in der Ebene in Klassen kongruenter oder in Klassen flächengleicher Dreiecke aufteilen; alle Funktionen von x kann man in Klassen einteilen, indem man alle jene Funktionen in einer Klasse zusammenfaßt, die in einem gegebenen Punkt gleiche Werte annehmen usw.

Die Kriterien, nach denen die Elemente einer Menge in Klassen eingeteilt werden, können vielfältigster Art sein. Jedoch kann kein solches Kriterium vollkommen willkürlich sein. Nehmen wir beispielsweise an, wir wollten die reellen Zahlen in Klassen einteilen, indem wir die Zahl b dann und nur dann derselben Klasse zuteilen wie die Zahl a, wenn $b > a$ ist. Es ist klar, daß man auf diesem Wege keine Klassen-

einteilung der reellen Zahlen erhalten kann, denn ist $b > a$ (d. h., daß b zu derselben Klasse wie a gehört), so ist $a < b$ (d. h., man kann a nicht derselben Klasse zuteilen wie b). *Außerdem kann a nicht in einer Klasse mit sich selbst zusammengefaßt werden,* da a nicht größer ist als a selbst! Ein anderes Beispiel: Wir wollen versuchen, die Punkte der Ebene in Klassen einzuteilen, indem wir zwei Punkte dann und nur dann der gleichen Klasse zuordnen, wenn die Entfernung zwischen ihnen kleiner als 1 ist. Es ist klar, daß man das nicht erreichen kann, denn wenn die Entfernung von a nach b kleiner als 1 und die Entfernung von b nach c kleiner als 1 ist, so heißt das noch lange nicht, daß die Entfernung von a nach c auch kleiner als 1 ist. Deshalb kann es passieren, daß dadurch, daß wir a derselben Klasse wie b und b derselben Klasse wie c zuteilen, zwei Punkte in dieselbe Klasse fallen, deren Entfernung voneinander größer als 1 ist.

Die angeführten Beispiele zeigen die Voraussetzungen, unter denen ein Kriterium tatsächlich eine Klasseneinteilung der Elemente einer Menge gestattet. Es sei M eine Menge, und gewisse Paare (a, b) von Elementen dieser Menge seien „ausgezeichnet".[1]) Ist (a, b) ein solches „ausgezeichnetes" Paar, so sagen wir, daß das Element a mit b durch die Relation φ verknüpft ist, und drücken das durch das Symbol $a \underset{\varphi}{\sim} b$ aus. Wenn wir beispielsweise die Einteilung der Dreiecke in Klassen flächengleicher Dreiecke betrachten, so bedeutet $a \underset{\varphi}{\sim} b$, „das Dreieck a hat den selben Flächeninhalt wie das Dreieck b". Eine gegebene Relation φ heißt *Äquivalenzrelation,* wenn sie folgende Eigenschaften besitzt:

1. Reflexivität: $a \underset{\varphi}{\sim} a$ für ein beliebiges Element $a \in M$.

2. Symmetrie: Ist $a \underset{\varphi}{\sim} b$, so ist auch $b \underset{\varphi}{\sim} a$.

3. Transitivität: Ist $a \underset{\varphi}{\sim} b$ und $b \underset{\varphi}{\sim} c$, so ist auch $a \underset{\varphi}{\sim} c$.

Diese Bedingungen sind notwendig und hinreichend dafür, daß die Relation φ (Kriterium!) eine Einteilung der Menge M in Klassen zuläßt. Jede Klasseneinteilung einer gegebenen Menge definiert auch eine gewisse Äquivalenzrelation. Denn eine Relation ist, wenn $a \underset{\varphi}{\sim} b$ bedeutet „a gehört zur selben Klasse wie b", reflexiv, symmetrisch und transitiv, wie leicht zu sehen ist.

Nun sei umgekehrt φ eine Äquivalenzrelation zwischen Elementen der Menge M und K_a die Klasse derjenigen Elemente x aus M, die einem gegebenen Element a äquivalent sind: $x \underset{\varphi}{\sim} a$. Auf Grund der Reflexivität gehört das Element a selbst der Klasse K_a an. Wir zeigen nun, daß zwei Klassen K_a und K_b entweder übereinstimmen oder disjunkt sind. Angenommen, ein Element c gehöre gleichzeitig zu K_a und K_b, d. h. $c \underset{\varphi}{\sim} a$ und $c \underset{\varphi}{\sim} b$. Dann gilt wegen der Symmetrie $a \underset{\varphi}{\sim} c$ und wegen der Transitivität

$$a \underset{\varphi}{\sim} b. \tag{1}$$

[1]) Hierbei wird die Reihenfolge der Elemente a und b berücksichtigt; (a, b) und (b, a) sind somit im allgemeinen verschiedene Paare.

Ist jetzt x ein beliebiges Element aus K_a, d. h. $x \underset{\varphi}{\sim} a$, so gilt wegen (1) und der Transitivität $x \underset{\varphi}{\sim} b$, d. h. $x \in K_b$. Genauso zeigt man, daß jedes Element $y \in K_b$ auch zu K_a gehört. Deshalb stimmen zwei Klassen K_a und K_b, die auch nur ein gemeinsames Element enthalten, überein. Wir haben damit eine Klasseneinteilung der Menge M durch eine vorgegebene Äquivalenzrelation erhalten.

Der Begriff der Klasseneinteilung einer Menge hängt eng mit dem in 1.2.1. besprochenen Abbildungsbegriff zusammen.

Es sei f eine Abbildung einer Menge A in eine Menge B. Wenn wir alle jene Elemente aus A, deren Bilder in B zusammenfallen, in einer Klasse zusammenfassen, so erhalten wir offenbar eine gewisse Klasseneinteilung der Menge A. Wir betrachten umgekehrt eine beliebige Menge A und irgendeine Einteilung von A in Klassen. B sei die Gesamtheit der Klassen, in die die Menge A eingeteilt ist. Indem wir jedem Element $a \in A$ diejenige Klasse (d. h. dasjenige Element aus B) zuordnen, zu der a gehört, so erhalten wir eine Abbildung der Menge A auf die Menge B.

Beispiele

1. Wir projizieren die x,y-Ebene auf die x-Achse. Die Urbilder der Punkte der x-Achse sind vertikale Geraden. Folglich entspricht dieser Abbildung eine Einteilung der Ebene in parallele Geraden.

2. Wir teilen alle Punkte des dreidimensionalen Raumes in Klassen ein, indem wir diejenigen Punkte in einer Klasse zusammenfassen, die die gleiche Entfernung vom Koordinatenursprung haben. Jede Klasse stellt eine Kugeloberfläche mit einem gewissen Radius dar. Die Gesamtheit aller dieser Klassen kann man mit der Menge aller Punkte identifizieren, die auf der Halbachse $[0, \infty)$ liegen. Folglich entspricht der Einteilung des dreidimensionalen Raumes in konzentrische Kugeloberflächen eine Abbildung eines Raumes auf eine Halbgerade.

3. Wir fassen alle reellen Zahlen mit gleichem gebrochenem Anteil in einer Klasse zusammen. Dieser Klasseneinteilung entspricht die Abbildung einer Geraden auf eine Kreislinie der Länge 1.

Der Äquivalenzbegriff ist ein Spezialfall des allgemeineren Begriffes der binären Relation. Es sei M eine beliebige Menge. Wir bezeichnen mit $M \times M$ oder M^2 die Gesamtheit aller geordneten Paare (a, b) mit $a, b \in M$. Man sagt, daß in M eine *binäre Relation* φ gegeben ist, wenn in M^2 eine gewisse Teilmenge R_φ ausgesondert wird. Genauer ausgedrückt sagen wir, daß a dann und nur dann in der Relation φ zum Element b steht — hierfür wird $a\varphi b$ geschrieben — wenn (a, b) zu R_φ gehört. Als Beispiel einer binären Relation möge die Identitätsrelation ε dienen. Es ist nämlich $a\varepsilon b$ dann und nur dann, wenn $a = b$ ist; dies ist eine Relation, die die Diagonale Δ in $M \times M$, d. h. die Teilmenge aller Paare der Gestalt (a, a) aussondert.

Es ist klar, daß jede Äquivalenzrelation in einer Menge M eine binäre Relation ist, die folgende Bedingungen erfüllt:

1. Die Diagonale Δ gehört zu R_φ (Reflexivität).

2. Ist $(a, b) \in R_\varphi$, so ist auch $(b, a) \in R_\varphi$ (Symmetrie).

3. Ist $(a, b) \in R_\varphi$ und $(b, c) \in R_\varphi$, so ist auch $(a, c) \in R_\varphi$ (Transitivität).

Folglich ist die Äquivalenzrelation eine binäre Relation, die die Bedingungen der Reflexivität, der Transitivität und der Symmetrie erfüllt.

In 1.4. werden wir einen anderen wichtigen Spezialfall einer binären Relation betrachten, die Halbordnung.

1.3. Äquivalenz von Mengen. Der Begriff der Mächtigkeit einer Menge

1.3.1. Endliche und unendliche Mengen. Bei einer konkreten Menge kann man manchmal, zumindest im Prinzip, die Anzahl ihrer Elemente angeben. So zum Beispiel bei der Menge aller Ecken eines Polyeders, der Menge aller Primzahlen, die eine gegebene Zahl nicht überschreiten, der Menge aller Wassermoleküle auf der Erde usw. Jede dieser Mengen enthält eine endliche, wenn auch möglicherweise uns unbekannte Anzahl von Elementen. Andererseits gibt es Mengen, die aus unendlich vielen Elementen bestehen, so beispielsweise die Menge aller natürlichen Zahlen, die Menge aller Punkte einer Geraden, aller Kreise in der Ebene, aller Polynome mit rationalen Koeffizienten usw. Damit, daß wir sagen, die Menge sei unendlich, wollen wir ausdrücken, daß man aus ihr ein Element, zwei Elemente usw. herausnehmen kann, wobei nach jedem derartigen Schritt in dieser Menge noch Elemente verbleiben.

Zwei endliche Mengen können wir bezüglich der Anzahl ihrer Elemente vergleichen und beurteilen, ob diese Anzahl gleich ist oder ob sie bei der einen Menge größer ist als bei der anderen. Nun ergibt sich die Frage, ob man unendliche Mengen auf ähnliche Art und Weise vergleichen kann, d. h., ob es Sinn hat, beispielsweise die Frage zu stellen, ob es mehr Kreise in der Ebene als rationale Punkte auf der Geraden oder Funktionen, die auf dem Intervall $[0, 1]$ definiert sind, oder Geraden im Raum usw. gibt.

Wir wollen untersuchen, wie sich zwei endliche Mengen vergleichen lassen. Man kann beispielsweise die Elemente beider Mengen abzählen und die Mengen auf diese Weise vergleichen. Man kann aber auch anders vorgehen, indem man nämlich versucht, eine *bijektive Abbildung* (*Bijektion*) herzustellen, d. h. eine umkehrbar eindeutige Zuordnung der Elemente dieser Mengen, mit anderen Worten eine Zuordnung, durch die jedem Element der einen Menge ein und nur ein Element der anderen Menge entspricht und umgekehrt. Es ist klar, daß sich eine umkehrbar eindeutige Zuordnung zwischen zwei endlichen Mengen dann und nur dann finden läßt, wenn die Anzahl ihrer Elemente gleich ist. Beispielsweise kann man, um zu prüfen, ob die Anzahl der Studenten einer Gruppe und die der Sitzplätze in einem Hörsaal übereinstimmen, ohne diese nachzuzählen, jeden Studenten auf einen bestimmten Platz setzen. Wenn die Plätze insgesamt ausreichen und kein einziger leerer Platz übrig-

bleibt, d. h., wenn eine bijektive Zuordnung zwischen diesen beiden Mengen hergestellt ist, so heißt das auch, daß die Anzahl ihrer Elemente übereinstimmt.

Ist das erste Verfahren (das Abzählen der Elemente) nur zum Vergleich endlicher Mengen geeignet, so ist das zweite Verfahren (die Herstellung einer eineindeutigen Zuordnung) auch für unendliche Mengen brauchbar.

1.3.2. Abzählbare Mengen. Die einfachste unter den unendlichen Mengen ist die Menge der natürlichen Zahlen. Jede Menge, deren Elemente man bijektiv allen natürlichen Zahlen zuordnen kann, bezeichnen wir als *abzählbare Menge*. Mit anderen Worten, eine abzählbare Menge ist eine solche Menge, deren Elemente man in einer unendlichen Folge durchnumerieren kann: $a_1, a_2, \ldots, a_n, \ldots$ Wir wollen einige Beispiele abzählbarer Mengen anführen.

1. *Die Menge aller ganzen Zahlen.* Wir konstruieren die Zuordnung zwischen allen ganzen und allen natürlichen Zahlen nach folgendem Schema:

$$\begin{array}{ccccc} 0 & -1 & 1 & -2 & 2\ldots, \\ 1 & 2 & 3 & 4 & 5\ldots \end{array}$$

Wir ordnen also allgemein der nichtnegativen Zahl $n \geqq 0$ die ungerade Zahl $2n + 1$ und der negativen Zahl $n < 0$ die gerade Zahl $2\,|n|$ zu:

$$n \leftrightarrow 2n + 1 \quad \text{für} \quad n \geqq 0,$$

$$n \leftrightarrow 2\,|n| \quad \text{für} \quad n < 0.$$

2. *Die Menge aller geraden positiven Zahlen.* Die Zuordnung ist offensichtlich:

$$n \leftrightarrow 2n.$$

3. *Die Menge* $2, 4, 8, \ldots, 2^n, \ldots$ *von Potenzen der Zahl* 2. Hier ist die Zuordnung auch klar: Jeder Zahl 2^n entspricht die Zahl n.

4. Wir betrachten nun ein kompliziertes Beispiel, und zwar wollen wir zeigen, daß *die Menge aller rationalen Zahlen abzählbar ist.* Jede rationale Zahl läßt sich eindeutig in Gestalt eines unkürzbaren Bruches $\alpha = \dfrac{p}{q}$, $q > 0$, darstellen. Wir nennen die Summe $|p| + q$ die *Höhe* der rationalen Zahl α. Es ist klar, daß die Anzahl der Brüche mit einer gegebenen Höhe n endlich ist. Die Höhe 1 beispielsweise hat nur die Zahl $\dfrac{0}{1}$, die Höhe 2 haben die Zahlen $\dfrac{1}{1}$ und $\dfrac{-1}{1}$, die Höhe 3 die Zahlen

$$\frac{2}{1}, \frac{1}{2}, \frac{-2}{1}, \frac{-1}{2},$$

usw. Wir numerieren nun alle rationalen Zahlen nach wachsender Höhe durch, d. h., zuerst schreiben wir die Zahlen der Höhe 1 auf, dann die Zahlen der Höhe 2

usw. Dadurch erhält jede rationale Zahl eine bestimmte Nummer, d. h., es ist eine eineindeutige Zuordnung zwischen allen natürlichen und allen rationalen Zahlen hergestellt.

Eine unendliche Menge, die nicht abzählbar ist, nennt man eine *überabzählbare* Menge.

Wir wollen einige allgemeine Eigenschaften abzählbarer Mengen angeben.

1. *Jede Teilmenge einer abzählbaren Menge ist endlich oder abzählbar.*

Beweis. Es sei A eine abzählbare Menge und B eine Teilmenge von A. Wir numerieren die Elemente der Menge A folgendermaßen: $a_1, a_2, \ldots, a_n, \ldots$

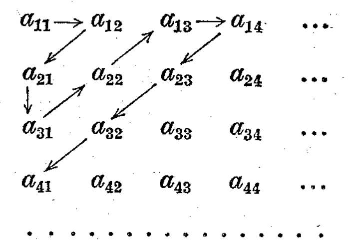
Diejenigen Elemente von A, die zu B gehören, bezeichnen wir mit $a_{n_1}, a_{n_2}, \ldots$ Wenn es unter den Zahlen $n_1, n_2, \ldots$ eine größte gibt, so ist B endlich, anderenfalls ist B abzählbar, da ihre Elemente $a_{n_1}, a_{n_2}, \ldots$ wiederum mit den Zahlen $1, 2, \ldots$ numeriert sind.

2. *Die Vereinigung einer endlichen oder abzählbaren Menge von abzählbaren Mengen ist ebenfalls eine abzählbare Menge.*

Beweis. Es seien $A_1, A_2, \ldots$ abzählbare Mengen. Wir können voraussetzen, daß sie paarweise disjunkt sind, da sonst die Mengen $A_1, A_2 \setminus A_1, A_3 \setminus (A_1 \cup A_2), \ldots$ (von denen wieder jede höchstens abzählbar ist und die dieselbe Vereinigung wie die Mengen $A_1, A_2, \ldots$ haben) betrachtet werden können. Alle Elemente der Mengen $A_1, A_2, \ldots$ kann man in Gestalt des folgenden unendlichen Schemas aufschreiben:

$$\begin{array}{ccccc} a_{11} & a_{12} & a_{13} & a_{14} & \ldots \\ a_{21} & a_{22} & a_{23} & a_{24} & \ldots \\ a_{31} & a_{32} & a_{33} & a_{34} & \ldots \\ a_{41} & a_{42} & a_{43} & a_{44} & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \end{array}$$

Dabei stehen in der ersten Zeile die Elemente der Menge A_1, in der zweiten die Elemente der Menge A_2 usw. Wir numerieren jetzt alle diese Elemente „diagonal", d. h., als erstes Element nehmen wir a_{11}, als zweites a_{12}, als drittes a_{21} usw., und zwar in der Richtung wachsend, in der die Pfeile in dem folgenden Schema weisen:



Es ist klar, daß dadurch jedes Element jeder der Mengen eine bestimmte Nummer erhält, d. h., es wird eine eineindeutige Zuordnung zwischen allen Elementen aller Mengen $A_1, A_2, \ldots$ und allen natürlichen Zahlen hergestellt. Unsere Behauptung ist damit bewiesen.

Aufgaben

1. Man zeige, daß die Menge aller Polynome mit rationalen Koeffizienten abzählbar ist.

2. Eine Zahl ξ nennt man *algebraisch*, wenn sie Nullstelle eines Polynoms mit rationalen Koeffizienten ist. Man zeige, daß die Menge aller algebraischen Zahlen abzählbar ist.

3. Man zeige, daß die Menge aller rationalen Integrale (d. h. der Integrale mit rationalen Grenzen) auf der Geraden abzählbar ist.

4. Man zeige, daß die Menge aller Punkte der Ebene mit rationalen Koordinaten abzählbar ist.

Hinweis. Man benutze die Eigenschaft 2.

3. *Jede unendliche Menge enthält eine abzählbare Teilmenge.*

Beweis. Es sei M eine unendliche Menge. Wir entnehmen ihr ein beliebiges Element a_1. Da M unendlich ist, findet sich in M ein Element a_2, das von a_1 verschieden ist; dazu gibt es ein Element a_3, das von a_1 und a_2 verschieden ist usw. Wenn wir diesen Prozeß (der nicht wegen „Mangels" an Elementen abbrechen kann, da M unendlich ist) fortsetzen, erhalten wir eine abzählbare Teilmenge

$$A = \{a_1, a_2, \ldots, a_n, \ldots\}$$

der Menge M. Damit ist die Aussage bewiesen.

Diese Aussage zeigt, daß unter den unendlichen Mengen die abzählbaren die „kleinsten" sind. Später werden wir klären, ob auch überabzählbare unendliche Mengen existieren.

1.3.3. Äquivalenz von Mengen. Dadurch, daß wir eine unendliche Menge mit der Menge der natürlichen Zahlen verglichen haben, sind wir zum Begriff der abzählbaren Mengen gekommen. Es ist klar, daß man Mengen nicht nur mit der Menge der natürlichen Zahlen vergleichen kann; der Begriff der umkehrbar eindeutigen Zuordnung (bijektiven Abbildung) gestattet den Vergleich zweier beliebiger Mengen miteinander. Wir wollen die folgende Definition einführen.

Definition. Zwei Mengen M und N heißen *äquivalent* (Bezeichnung $M \sim N$), wenn sich zwischen ihren Elementen eine eineindeutige Zuordnung herstellen läßt.

Der Begriff der Äquivalenz ist anwendbar auf beliebige Mengen, sowohl endliche als auch unendliche. Zwei endliche Mengen sind dann (und nur dann) äquivalent, wenn die Anzahl ihrer Elemente übereinstimmt. Die Definition der abzählbaren Menge läßt sich jetzt folgendermaßen formulieren: *Eine Menge heißt abzählbar, wenn sie der Menge der natürlichen Zahlen äquivalent ist.*

Es ist klar, daß zwei Mengen, die einer dritten äquivalent sind, auch zueinander äquivalent sind; insbesondere sind alle abzählbaren Mengen zueinander äquivalent.

Beispiele

1. Die Mengen der Punkte zweier beliebiger Strecken $[a, b]$ und $[c, d]$ sind äquivalent. Aus Abb. 5 wird klar, wie man zwischen ihnen eine bijektive Abbildung herstellen kann. Die Punkte p und q entsprechen einander genau dann, wenn sie auf ein und demselben Strahl liegen, der aus dem Punkt O kommt, in dem sich die Geraden ac und bd schneiden.

2. Die Menge aller Punkte auf der erweiterten komplexen Ebene ist der Menge aller Punkte auf der Kugeloberfläche äquivalent. Die bijektive Abbildung $\alpha \leftrightarrow z$ kann man beispielsweise mit Hilfe der stereographischen Projektion erhalten (Abb. 6).

3. Die Menge aller Zahlen im Intervall (0, 1) ist der Menge aller Punkte auf der Geraden äquivalent. Eine Zuordnung kann man beispielsweise mit Hilfe der Funktion

$$y = \frac{1}{\pi} \arctan x + \frac{1}{2}$$

herstellen.

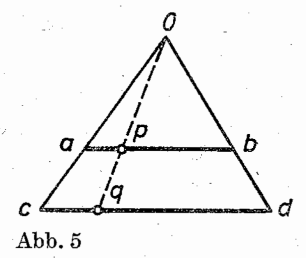


Abb. 5

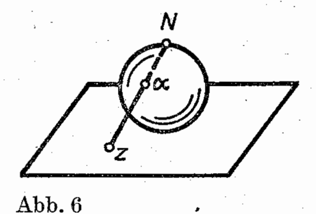


Abb. 6

Betrachtet man die Beispiele, die hier und in 1.3.2. angeführt wurden, so kann man feststellen, daß sich eine unendliche Menge manchmal als äquivalent zu einer ihrer echten Teilmengen erweist. Beispielsweise zeigt es sich, daß es „ebensoviele" natürliche Zahlen wie ganze Zahlen oder rationale Zahlen gibt; auf dem Intervall (0, 1) gibt es „ebensoviele" Punkte wie auf der ganzen Geraden usw. Dieses Phänomen ist charakteristisch für unendliche Mengen. Wir haben nämlich in 1.3.2. (Eigenschaft 3) gezeigt, daß man aus jeder unendlichen Menge eine abzählbare Teilmenge aussondern kann. Es sei $A = \{a_1, a_2, \ldots, a_n, \ldots\}$ eine solche Teilmenge. Wir wollen sie in zwei abzählbare Teilmengen

$$A_1 = \{a_1, a_3, a_5, \ldots\} \quad \text{und} \quad A_2 = \{a_2, a_4, a_6, \ldots\}$$

aufteilen und zwischen A und A_1 eine eineindeutige Zuordnung herstellen. Diese Zuordnung läßt sich zu einer eineindeutigen Zuordnung zwischen den Mengen

$A \cup (M \setminus A) = M$ und $A_1 \cup (M \setminus A) = M \setminus A_2$ erweitern, indem man jedes Elemente aus $M \setminus A$ sich selbst zuordnet. Außerdem fällt die Menge $M \setminus A_2$ nicht mit M zusammen, d. h., sie ist eine echte Teilmenge von M. Auf diese Weise erhalten wir folgenden Satz:

Jede unendliche Menge ist einer ihrer echten Teilmengen äquivalent.

Diese Eigenschaft kann man als Definition der unendlichen Menge auffassen.

Aufgabe. Man zeige $M \sim M \cup A$, wenn M eine beliebige unendliche Menge und A eine abzählbare Menge ist.

1.3.4. Die Überabzählbarkeit der Menge der reellen Zahlen. In 1.3.2. haben wir Beispiele abzählbarer Mengen angegeben. Die Zahl dieser Beispiele ließe sich noch vergrößern. Außerdem ist, wie wir gezeigt haben, die Vereinigung endlich vieler oder abzählbar vieler abzählbarer Mengen ebenfalls eine abzählbare Menge.

Natürlich stellt sich nun die Frage: Gibt es überhaupt überabzählbare Mengen? Eine positive Antwort darauf gibt der folgende Satz.

Satz 1. *Die Menge der reellen Zahlen, die zwischen* 0 *und* 1 *liegen, ist überabzählbar.*

Beweis. Wir nehmen an, daß irgendeine abzählbare Menge (aller oder nur einiger) reeller Zahlen α, die in dem Intervall $[0, 1]$ liegen, gegeben ist:

$$\left.\begin{array}{l} \alpha_1 = 0, a_{11}a_{12}a_{13} \ldots a_{1n} \ldots, \\ \alpha_2 = 0, a_{21}a_{22}a_{23} \ldots a_{2n} \ldots, \\ \alpha_3 = 0, a_{31}a_{32}a_{33} \ldots a_{34} \ldots, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_n = 0, a_{n1}a_{n2}a_{n3} \ldots a_{nn} \ldots, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{array}\right\} \qquad (1)$$

Hierbei ist a_{ik} die k-te Dezimalziffer der Zahl α_i. Wir konstruieren nun nach dem *Cantorschen Diagonalverfahren* den Bruch

$$\beta = 0, b_1b_2b_3 \ldots b_n \ldots$$

folgendermaßen: Für b_1 wählen wir eine beliebige Ziffer, die nicht mit a_{11} übereinstimmt, für b_2 eine beliebige Ziffer, die nicht mit a_{22} übereinstimmt usw., allgemein wählen wir für b_n eine beliebige Ziffer, die nicht mit a_{nn} übereinstimmt. Dieser Dezimalbruch kann nicht mit einem Bruch übereinstimmen, der in (1) enthalten ist. Er unterscheidet sich nämlich von α_1 mindestens in der ersten Ziffer, von α_2 in der zweiten Ziffer usw.; im allgemeinen ist der Bruch β wegen $b_n \neq a_{nn}$ für alle n von jedem Bruch α_i verschieden, der zu (1) gehört. Aus diesem Grunde füllt keine abzählbare Menge reeller Zahlen, die in den Intervall $[0, 1]$ liegen, dieses Intervall aus.

Der hier geführte Beweis enthält eine kleine Ungenauigkeit, denn gewisse Zahlen (nämlich Zahlen der Gestalt $p/10^q$) können auf zweierlei Art als Dezimalbruch geschrieben werden: mit unendlich vielen Nullen oder mit unendlich vielen Neunen. Zum Beispiel ist

$$\frac{1}{2} = \frac{5}{10} = 0{,}5000\ldots = 0{,}4999\ldots$$

Deshalb garantiert ein Nichtübereinstimmen zweier Dezimalbrüche noch nicht die Verschiedenheit der durch sie dargestellten Zahlen. Wenn wir jedoch den Bruch β sorgfältiger konstruieren, so daß er weder Nullen noch Neunen enthält, indem wir beispielsweise $b_n = 2$ setzen, wenn $a_{nn} = 1$ ist, und $b_n = 1$, wenn $a_{nn} \neq 1$ ist, so wird der Beweis vollständig korrekt.

Aufgabe. Man zeige, daß die Zahlen, die zwei verschiedene Dezimaldarstellungen besitzen, eine abzählbare Menge bilden.

Das Intervall [0, 1] gibt uns also ein Beispiel einer überabzählbaren Menge. Wir beschreiben nun einige Beispiele von Mengen, die dem Intervall [0, 1] äquivalent sind:

1. Die Menge aller Punkte einer beliebigen Strecke $[a, b]$ oder eines Intervalls (a, b).

2. Die Menge aller Punkte auf der Geraden.

3. Die Menge aller Punkte der Ebene, des Raumes, der Kugeloberfläche, und die Menge der Punkte, die innerhalb der Kugel liegen, usw.

4. Die Menge aller Geraden in der Ebene.

5. Die Menge aller stetigen Funktionen einer oder mehrerer Variablen.

In den Fällen 1 und 2 braucht kein Beweis mehr geführt zu werden (man vergleiche mit den Beispielen 1 und 3 aus 1.3.3.). In den übrigen Fällen ist ein direkter Beweis kompliziert.

Aufgabe. Unter Benutzung der obigen Resultate und der Aufgabe 2 aus 1.3.2. beweise man die Existenz von *transzendenten* Zahlen, d. h. von Zahlen, die nicht algebraisch sind.

1.3.5. Der Satz von Cantor-Bernstein. Der folgende Satz ist einer der grundlegenden Sätze der Mengenlehre.

Satz 2 (Cantor-Bernstein). *A und B seien zwei beliebige Mengen. Existiert eine eineindeutige Abbildung f der Menge A auf eine Teilmenge B_1 der Menge B und existiert eine eineindeutige Abbildung g der Menge B auf eine Teilmenge A_1 der Menge A, so sind A und B äquivalent.*

Beweis[1]). Ohne Beschränkung der Allgemeinheit können wir annehmen, daß A und B disjunkt sind. Ein Element x aus A oder aus B heißt *Vorgänger* eines Elementes y aus A, wenn man y aus x durch einmalige oder mehrmalige Anwendung der Abbildungen f und g gewinnen kann. Wir teilen jetzt die Elemente aus A auf folgende Weise in disjunkte Mengen ein: A_E ist die Gesamtheit der Elemente mit einer geradzahligen Anzahl von Vorgängern (hierzu gehören auch die Elemente aus $A \setminus A_1$, die keinen Vorgänger haben), A_O ist die Gesamtheit der Elemente mit einer ungeradzahligen Anzahl von Vorgängern, und A_I ist die Gesamtheit der Elemente mit unendlich vielen Vorgängern. (Hierbei kann die Anzahl der verschiedenen Vorgänger endlich sein. Dieser Fall tritt ein, wenn das obige Verfahren der Konstruktion der Vorgänger zyklisch ist, wobei ein Zyklus nur endlich viele verschiedene Elemente enthält.) Entsprechend wird die Menge B unterteilt. Wir bemerken jetzt, daß f die Menge A_E auf B_O und die Menge A_I auf B_I abbildet, während g^{-1} die Menge A_O auf B_E abbildet. Folglich ist die Abbildung ψ, die auf $A_E \cup A_I$ mit f und auf A_O mit g^{-1} übereinstimmt, eine eineindeutige Abbildung von ganz A auf ganz B. Damit ist der Satz bewiesen.

1.3.6. Der Begriff der Mächtigkeit einer Menge. Wenn zwei endliche Mengen äquivalent sind, bestehen sie aus der gleichen Anzahl von Elementen. Sind die äquivalenten Mengen M und N beliebig, so sagt man, M und N haben die gleiche *Mächtigkeit*. Damit ist die Mächtigkeit etwas, was alle einander äquivalenten Mengen gemeinsam haben. Für endliche Mengen stimmt der Begriff der Mächtigkeit mit dem gewöhnlichen Begriff der Anzahl der Elemente einer Menge überein. Die Mächtigkeit der Menge der natürlichen Zahlen (d. h. einer beliebigen abzählbaren Menge) bezeichnet man mit dem Symbol $\aleph_0$ (gelesen: Aleph null). Die Mächtigkeit von Mengen, die der Menge der reellen Zahlen des Intervalls $[0, 1]$ äquivalent sind, nennt man die *Mächtigkeit des Kontinuums*. Diese Mächtigkeit bezeichnet man mit dem Symbol c (oder dem Symbol $\aleph$).

Die sehr tiefliegende Frage der Existenz von Mächtigkeiten, die zwischen $\aleph_0$ und c liegen, wird weiter unten in 1.4. berührt. Im allgemeinen sind unendliche Mengen, die in der Analysis auftreten, entweder abzählbar, oder sie haben die Mächtigkeit des Kontinuums.

Für Mächtigkeiten endlicher Mengen, d. h. für natürliche Zahlen, haben wir außer dem Begriff der Gleichheit auch die Begriffe „größer als" und „kleiner als". Wir wollen sehen, wie man diese Begriffe auf unendliche Mengen erweitern kann.

Es seien A und B zwei beliebige Mengen mit ihren Mächtigkeiten $m(A)$ und $m(B)$. Dann sind folgende Fälle möglich:

1. A ist einer Teilmenge von B äquivalent, und B ist einer Teilmenge von A äquivalent.

[1]) Anm. d. Übers.: Der Beweis wurde bei der Übersetzung geringfügig abgeändert.

2. A enthält eine Teilmenge, die zu B äquivalent ist, aber in B gibt es keine zu A äquivalente Teilmenge.

3. B enthält eine Teilmenge, die zu A äquivalent ist, aber in A gibt es keine zu B äquivalente Teilmenge.

4. In keiner der beiden Mengen gibt es Teilmengen, die zu der anderen Menge äquivalent sind.

Im ersten Fall sind die Mengen auf Grund des Satzes von CANTOR-BERNSTEIN einander äquivalent, d. h.

$$m(A) = m(B).$$

Im zweiten Fall wird man $m(A) > m(B)$ und im dritten Fall $m(A) < m(B)$ schreiben.

Im vierten Fall schließlich müßten wir folgern, daß die Mächtigkeiten der Mengen A und B nicht miteinander vergleichbar sind. Dieser Fall ist nicht möglich! Das folgt aus dem Satz von ZERMELO, von dem in 1.4. die Rede sein wird. *Folglich sind zwei beliebige Mengen A und B entweder einander äquivalent* (und dann gilt $m(A) = m(B)$), *oder sie genügen einer der beiden Relationen*

$$m(A) < m(B) \quad \textit{oder} \quad m(A) > m(B).$$

Wir haben oben die Bemerkung gemacht, daß abzählbare Mengen die „kleinsten" unter den unendlichen Mengen sind, und dann gezeigt, daß es unendliche Mengen gibt, deren Unendlichkeit eine „höhere Ordnung" hat, — das sind die Mengen von der Mächtigkeit des Kontinuums. Existieren unendliche Mengen, die die Mächtigkeit des Kontinuums übertreffen? Allgemein ausgedrückt, existiert irgendeine „höhere" Mächtigkeit oder nicht? Es zeigt sich, daß folgender Satz gilt.

Satz 3. *Es sei M irgendeine Menge und $\mathfrak{M}$ eine Menge, deren Elemente alle möglichen Teilmengen der Menge M sind. Dann hat $\mathfrak{M}$ eine höhere Mächtigkeit als die Ausgangsmenge M.*[1])

Beweis. Es ist leicht zu sehen, daß die Mächtigkeit $\mathfrak{m}$ der Menge $\mathfrak{M}$ nicht kleiner sein kann als die Mächtigkeit m der Ausgangsmenge M, denn die „einelementigen" Teilmengen von M bilden in $\mathfrak{M}$ eine Teilmenge, die der Menge M äquivalent ist. Es bleibt zu zeigen, daß die Mächtigkeiten $\mathfrak{m}$ und m nicht gleich sind. Zwischen den Elementen $a, b, \ldots$ der Menge M und irgendwelchen Elementen $A, B, \ldots$ der Menge $\mathfrak{M}$ (d. h. irgendwelchen Teilmengen von M) sei eine eineindeutige Zuordnung hergestellt:

$$a \leftrightarrow A, b \leftrightarrow B, \ldots$$

Wir zeigen, daß diese mit Sicherheit nicht ganz $\mathfrak{M}$ ausschöpfen kann. Hierzu konstruieren wir eine solche Menge $X \subset M$, der kein Element aus M entspricht.

[1]) Anm. d. Übers.: $\mathfrak{M}$ wird auch *Potenzmenge* von M genannt.

Es sei X die Gesamtheit der Elemente aus M, die nicht zu denjenigen Teilmengen gehören, die ihnen entsprechen. Ausführlicher: Gilt $a \leftrightarrow A$ und $a \in A$, so zählen wir das Element a nicht zu X; gilt aber $a \leftrightarrow A$ und $a \notin A$, so zählen wir das Element a zu X. Es ist klar, daß X eine Teilmenge der Menge M, d. h. ein bestimmtes Element aus $\mathfrak{M}$ ist. Wir zeigen nun, daß der Menge X kein Element aus M entsprechen kann. Angenommen, es existiert ein solches Element $x \leftrightarrow X$, so wollen wir untersuchen, ob es zu X gehört oder nicht. Es sei $x \notin X$. Da aber nach Definition jedes Element zu X gehört, das nicht in der Menge enthalten ist, die ihm entspricht, muß das Element x zu X gehören. Wenn wir umgekehrt annehmen, daß x in X enthalten ist, finden wir, daß x nicht in X enthalten sein kann, da in X nur diejenigen Elemente enthalten sind, die nicht zu den ihnen entsprechenden Teilmengen gehören. Folglich muß das Element, das der Teilmenge X entspricht, gleichzeitig in X enthalten und nicht enthalten sein. Daraus folgt, daß es ein solches Element überhaupt nicht gibt, d. h., daß es unmöglich ist, eine eineindeutige Zuordnung zwischen den Elementen der Menge M und allen ihren Teilmengen herzustellen. Damit ist der Satz bewiesen.

Wir können also zu einer Menge beliebiger Mächtigkeit eine Menge noch höherer Mächtigkeit konstruieren, danach eine noch größere usw. Auf diese Weise erhalten wir eine nach oben unbegrenzte Skala von Mächtigkeiten.

Anmerkung. Die Mächtigkeit der Menge $\mathfrak{M}$ bezeichnet man mit dem Symbol 2^m, wobei m die Mächtigkeit von M ist. (Der Leser wird leicht den Sinn dieser Bezeichnungsweise verstehen, wenn er den Fall einer endlichen Menge M betrachtet.) Auf diese Weise kann man den vorangegangenen Satz durch die Ungleichung $m < 2^m$ ausdrücken. Insbesondere erhalten wir für $m = \aleph_0$ die Ungleichung $\aleph_0 < 2^{\aleph_0}$. Wir wollen zeigen, daß $2^{\aleph_0} = \aleph$ ist, d. h., wir zeigen, daß *die Mächtigkeit der Menge aller Teilmengen der Menge der natürlichen Zahlen gleich der Mächtigkeit des Kontinuums ist.*

Wir teilen die Teilmengen der Menge der natürlichen Zahlen in zwei Klassen $\mathfrak{P}$ und $\mathfrak{Q}$ ein, nämlich in diejenigen (Klasse $\mathfrak{P}$), deren Komplement unendlich ist, und in diejenigen (Klasse $\mathfrak{Q}$), deren Komplement endlich ist. Zur Klasse $\mathfrak{Q}$ gehört insbesondere die Menge der natürlichen Zahlen, da ihr Komplement leer ist. Die Anzahl der Teilmengen aus der Klasse $\mathfrak{Q}$ ist abzählbar. (Der Beweis sei dem Leser überlassen.) Die Klasse $\mathfrak{Q}$ hat keinen Einfluß auf die Mächtigkeit der Menge $\mathfrak{M} = \mathfrak{P} \cup \mathfrak{Q}$.

Zwischen den Teilmengen der Klasse $\mathfrak{P}$ und den reellen Zahlen des halboffenen Intervalls $[0, 1]$ läßt sich eine eineindeutige Zuordnung herstellen.

Wir ordnen der Teilmenge $A \in \mathfrak{P}$ in eineindeutiger Weise die Zahl α, $0 \leqq \alpha < 1$, mit der dyadischen Entwicklung

$$\alpha = \frac{\varepsilon_1}{2} + \frac{\varepsilon_2}{2^2} + \cdots + \frac{\varepsilon_n}{2^n} + \cdots$$

zu, wobei $\varepsilon_n = 1$ oder 0 ist, je nachdem, ob n zur Menge A gehört oder nicht. Das Nachprüfen der Details überlassen wir dem Leser.

Aufgabe. Man zeige, daß die Gesamtheit aller auf einer Menge M definierten numerischen Funktionen (oder allgemeiner derjenigen Funktionen, die Werte aus einer Menge annehmen, welche nicht weniger als zwei Elemente enthält) eine größere Mächtigkeit hat als die Menge M selbst.

Hinweis. Man bediene sich der Tatsache, daß die Menge aller *Indikatoren*, d. h. derjenigen Funktionen auf M, die nur die beiden Werte 0 und 1 annehmen, der Menge aller Teilmengen von M äquivalent ist.

1.4. Geordnete Mengen. Transfinite Zahlen

In diesem Abschnitt untersuchen wir eine Reihe von Begriffen, die mit dem Begriff der Ordnung von Mengen zusammenhängen. Wir beschränken uns hier auf eine erste Einführung; eine ausführlichere Darstellung kann man in der am Ende des Buches angegebenen Literatur finden.

1.4.1. Halbgeordnete Mengen. Es sei M eine beliebige Menge und φ eine binäre Relation in ihr (die durch eine Menge $R_\varphi \subset M \times M$ definiert ist). Wir nennen diese Relation eine *Halbordnung*, wenn sie folgende Bedingungen erfüllt:

1. *Reflexivität:* $a\varphi a$,
2. *Transitivität:* Ist $a\varphi b$ und $b\varphi c$, so gilt $a\varphi c$,
3. *Antisymmetrie:* Ist $a\varphi b$ und $b\varphi a$, so gilt $a = b$.

Eine Halbordnung wird üblicherweise mit dem Symbol $\leqq$ bezeichnet. Damit bedeutet $a \leqq b$, daß das Paar (a, b) der entsprechenden Menge R_φ angehört. Vom Element a sagt man in diesem Zusammenhang auch, daß es *nicht größer ist als* b oder daß es *unterhalb von* b liegt. Eine Menge, in der eine Halbordnung gegeben ist, nennt man *halbgeordnet*.

Wir wollen einige Beispiele halbgeordneter Mengen angeben.

1. Jede Menge kann man trivialerweise als halbgeordnet betrachten, wenn man genau dann $a \leqq b$ setzt, wenn $a = b$ ist. Anders ausgedrückt, als Halbordnung kann man immer die binäre Relation der Identität ε annehmen. Dieses Beispiel ist natürlich nicht von großem Interesse.

2. Es sei M die Menge aller stetigen Funktionen auf dem Intervall $[\alpha, \beta]$. Setzen wir $f \leqq g$ genau dann, wenn $f(t) \leqq g(t)$ für alle $t, \alpha \leqq t \leqq \beta$ gilt, so erhalten wir offenbar eine Halbordnung.

3. Die Menge aller Teilmengen einer fixierten Menge ist durch die Inklusion halbgeordnet: $M_1 \leqq M_2$ bedeutet, daß $M_1 \subset M_2$ gilt.

4. Die Menge aller natürlichen Zahlen ist halbgeordnet, wenn $a \leqq b$ bedeutet, daß „b ohne Rest durch a teilbar ist".

Es sei M eine beliebige halbgeordnete Menge. In dem Fall, daß $a \leqq b$ und $a \neq b$ ist, werden wir das Symbol $<$ benutzen, d. h., wir schreiben $a < b$ und sagen, daß

a kleiner als b ist oder daß *a streng unterhalb von b* liegt. Neben der Schreibweise $a \leqq b$ werden wir die gleichwertige Schreibweise $b \geqq a$ benutzen und damit sagen, daß *b nicht kleiner als a* (*größer als a*, wenn $b \neq a$) ist oder daß *b auf a folgt*. Ein Element a nennt man *maximal*, wenn aus $a \leqq b$ folgt, daß $a = b$ ist. Ein Element a nennt man *minimal*, wenn aus $c \leqq a$ folgt, daß $c = a$ ist.

Eine halbgeordnete Menge, in der sich für je zwei beliebige Punkte a, b ein ihnen folgender Punkt c finden läßt ($a \leqq c$, $b \leqq c$), nennt man *gerichtet*.

1.4.2. Ordnungstreue Abbildungen. Es seien M und M' zwei halbgeordnete Mengen, und f sei eine Abbildung von M in M'. Wir sagen, daß diese Abbildung *ordnungstreu* ist, wenn aus $a \leqq b$ mit $a, b \in M$ folgt, daß $f(a) \leqq f(b)$ (in M') ist. Die Abbildung f bezeichnet man als *Isomorphismus* der halbgeordneten Mengen M und M', wenn sie bijektiv ist und die Relation $f(a) \leqq f(b)$ dann und nur dann erfüllt ist, wenn $a \leqq b$ ist. Die Mengen M und M' nennt man dann (zueinander) *isomorph*.

Beispielsweise sei M die Menge der natürlichen Zahlen, die durch die „Teilbarkeit" halbgeordnet ist (siehe 1.4.1., Beispiel 4), und M' dieselbe Menge, aber im natürlichen Sinne geordnet, d. h., daß $b \geqq a$ genau dann gilt, wenn $b - a$ eine positive Zahl ist. Dann ist die Abbildung von M auf M', die jede Zahl n sich selbst zuordnet, ordnungstreu (jedoch kein Isomorphismus).

Die Isomorphismen zwischen halbgeordneten Mengen stellen eine Äquivalenzrelation dar. Folglich können wir, wenn wir irgendein System[1]) halbgeordneter Mengen haben, diese Mengen in Klassen einander isomorpher Mengen aufteilen. Es ist klar, daß man, wenn nicht der Charakter der Elemente einer Menge interessiert, sondern nur die in den Mengen enthaltene Halbordnung, zwei einander isomorphe halbgeordnete Mengen identifizieren kann.

1.4.3. Ordnungstypen. Geordnete Mengen. Von isomorphen halbgeordneten Mengen wollen wir sagen, daß sie den gleichen *Ordnungstyp* haben. Damit ist der Ordnungstyp für alle einander isomorphen halbgeordneten Mengen charakteristisch, ähnlich wie die Mächtigkeit für alle einander äquivalenten Mengen charakteristisch ist (unabhängig davon, ob in ihnen eine Ordnungsrelation existiert oder nicht).

Es seien a und b Elemente einer halbgeordneten Menge. Es kann sich herausstellen, daß weder $a \leqq b$ noch $b \leqq a$ gilt. In diesem Fall nennt man die Elemente a und b *unvergleichbar*. Damit ist die Ordnungsrelation nur für gewisse Paare von Elementen definiert, weshalb wir auch von einer Halbordnung sprechen. Wenn es jedoch in einer halbgeordneten Menge M keine unvergleichbaren Elemente gibt, so nennt man die Menge M *geordnet* (*linear geordnet, vollständig geordnet*). Folglich ist eine Menge M geordnet, wenn sie halbgeordnet ist und wenn für zwei beliebige voneinander verschiedene Elemente $a, b \in M$ entweder $a < b$ oder $b < a$ gilt.

---

[1]) Wir vermeiden Begriffe wie „alle halbgeordneten Mengen", da sie ähnlich dem Begriff „Menge aller Mengen" innerliche Widersprüche aufweisen und nicht in exakte mathematische Konzeptionen aufgenommen werden können.

Die Mengen aus 1.4.1., Beispiel 1 bis 4, sind nur halbgeordnet. Als einfachste Beispiele linear geordneter Mengen mögen die natürlichen Zahlen, die rationalen Zahlen, die reellen Zahlen auf dem Intervall [0, 1] usw. dienen (mit den natürlichen Relationen „größer als" und „kleiner als", die in diesen Mengen gegeben sind).

Es ist klar, daß jede Teilmenge einer geordneten Menge selbst geordnet ist. Da die Ordnung ein Spezialfall der Halbordnung ist, sind auf geordnete Mengen der Begriff der ordnungstreuen Abbildung und insbesondere der Begriff des Isomorphismus anwendbar. Deshalb kann man von einem Ordnungstyp der geordneten Menge sprechen.

Die Menge der natürlichen Zahlen 1, 2, 3, ... mit der natürlichen Ordnungsrelation zwischen ihren Elementen ist das einfachste Beispiel einer unendlichen geordneten Menge. Ihr Ordnungstyp wird normalerweise mit dem Symbol ω bezeichnet.

Wenn zwei halbgeordnete Mengen isomorph sind, haben sie natürlich die gleiche Mächtigkeit (der Isomorphismus ist insbesondere eine bijektive Abbildung), und deshalb kann man von der Mächtigkeit sprechen, die einem gegebenen Ordnungstyp entspricht (beispielsweise entspricht dem Typ ω die Mächtigkeit $\aleph_0$). Die Umkehrung ist allerdings falsch: Eine Menge einer gegebenen Mächtigkeit kann im allgemeinen auf verschiedene Arten geordnet sein. Nur der Ordnungstyp einer linear geordneten e n d l i c h e n Menge wird eindeutig durch die Zahl n ihrer Elemente definiert (und wird auch mit n bezeichnet). Schon für die abzählbare Menge der natürlichen Zahlen ist beispielsweise neben ihrem „natürlichen" Typ ω folgender Typ möglich:

$$1, 3, 5, \ldots, 2, 4, 6, \ldots,$$

d. h. ein Typ, bei dem eine beliebige gerade Zahl auf eine beliebige ungerade Zahl folgt, aber die ungeraden und geraden Zahlen untereinander der Größe nach geordnet sind. Man kann zeigen, daß die Anzahl der verschiedenen Ordnungstypen, die der Mächtigkeit $\aleph_0$ entsprechen, unendlich und sogar überabzählbar ist.

1.4.4. Die geordnete Vereinigung geordneter Mengen. Es seien M_1 und M_2 zwei disjunkte geordnete Mengen mit den Ordnungstypen θ_1 und θ_2. In der Vereinigung $M_1 \cup M_2$ der Mengen M_1 und M_2 läßt sich eine Ordnung einführen, indem wir annehmen, daß zwei Elemente aus M_1 geordnet sind wie in M_1, daß zwei Elemente aus M_2 geordnet sind wie in M_2 und daß jedes Element aus M_1 jedem Element aus M_2 vorangeht. (Man zeige, daß das tatsächlich eine lineare Ordnung ist.) Eine derartige geordnete Menge wollen wir die *geordnete Vereinigung* der Mengen M_1 und M_2 nennen und $M_1 + M_2$ schreiben. Wir betonen, daß hier die Reihenfolge der Summanden wichtig ist: Die Vereinigung $M_2 + M_1$ ist im allgemeinen nicht isomorph der Vereinigung $M_1 + M_2$.

Den Ordnungstyp der Vereinigung $M_1 + M_2$ wollen wir die *geordnete Summe* der Ordnungstypen θ_1 und θ_2 nennen und $\theta_1 + \theta_2$ schreiben.

Diese Definition läßt sich leicht auf eine beliebige endliche Anzahl von Summanden $\theta_1, \theta_2, \ldots, \theta_m$ erweitern.

Beispiel. Wir betrachten die Ordnungstypen ω und n. Es ist leicht zu sehen, daß $n + \omega = \omega$ ist; fügen wir nämlich der Menge der natürlichen Zahlen $1, 2, 3, \ldots, k, \ldots$ links eine endliche Anzahl von Zahlen hinzu, so erhalten wir wieder den Ordnungstyp ω. Andererseits ist der Ordnungstyp $\omega + n$, d. h. der Ordnungstyp der Menge

$$1, 2, 3, \ldots, k, \ldots, a_1, a_2, \ldots, a_n$$

offensichtlich nicht gleich ω.

1.4.5. Wohlgeordnete Mengen. Transfinite Zahlen. Wir haben oben die Begriffe der Halbordnung und der Ordnung eingeführt. Wir führen noch den schärferen sehr wichtigen Begriff der *Wohlordnung* ein.

Definition. Eine geordnete Menge heißt *wohlgeordnet*, wenn jede ihrer nichtleeren Teilmengen ein kleinstes (d. h. ein allen Elementen dieser Teilmenge vorangehendes) Element enthält.

Ist eine geordnete Menge endlich, so ist sie offensichtlich auch wohlgeordnet. Als Beispiel einer geordneten, aber nicht wohlgeordneten Menge möge das Intervall [0, 1] dienen. Die Menge selbst enthält ein kleinstes Element, die Zahl 0, aber die Teilmenge, die aus den positiven Zahlen besteht, enthält kein kleinstes Element.

Es ist klar, daß *jede (nicht leere) Teilmenge einer wohlgeordneten Menge selbst wohlgeordnet ist.*

Den Ordnungstyp einer wohlgeordneten Menge nennt man *Ordinalzahl* oder *Ordnungszahl* (*transfinite Ordnungszahl*, wenn man hervorheben will, daß von einer unendlichen Menge die Rede ist).

Die Menge der natürlichen Zahlen (mit der natürlichen Ordnungsrelation) ist nicht nur geordnet, sondern auch wohlgeordnet. Deshalb ist ihr Ordnungstyp ω eine (transfinite!) Ordnungszahl. Eine Ordnungszahl ist auch $\omega + k$, d. h. der Typ der Menge

$$1, 2, \ldots, n, \ldots, a_1, a_2, \ldots, a_k.$$

Dagegen ist die Menge

$$\ldots, -n, \ldots, -3, -2, -1 \tag{1}$$

zwar geordnet, jedoch nicht wohlgeordnet. Hier existiert in jeder nichtleeren Teilmenge ein größtes Element (d. h. eines, das auf alle anderen folgt), aber nicht immer ein kleinstes. (Beispielsweise besitzt die gesamte Menge (1) kein kleinstes Element.) Der Ordnungstyp der Menge (1) (der keine Ordnungszahl ist!) wird üblicherweise mit dem Symbol ω^* bezeichnet.

Wir beweisen die folgende einfache, aber wichtige Tatsache.

Lemma 1. *Die geordnete Vereinigung einer endlichen Anzahl wohlgeordneter Mengen ist eine wohlgeordnete Menge.*

Es sei M eine beliebige Teilmenge der geordneten Vereinigung $M_1 + M_2 + \cdots + M_n$ wohlgeordneter Mengen; wir betrachten die erste der Mengen M_k, die Elemente aus M enthält. Der Teil der Menge M, der zu M_k gehört, ist eine Teilmenge der wohlgeordneten Menge M_k und hat folglich ein erstes Element. Dieses Element ist auch das erste Element von ganz M.

Folgerung. *Die geordnete Summe von Ordnungszahlen ist ebenfalls eine Ordnungszahl.*

Wir können auf diese Weise, indem wir von einem gewissen Vorrat von Ordnungszahlen ausgehen, neue Ordnungszahlen konstruieren. Wenn man beispielsweise von natürlichen Zahlen (d. h. von endlichen Ordnungszahlen) und der Ordnungszahl ω ausgeht, so kann man die Ordnungszahlen

$$\omega + n, \ \omega + \omega, \ \omega + \omega + n, \ \omega + \omega + \omega \text{ usw.}$$

erhalten.

Der Leser wird leicht wohlgeordnete Mengen konstruieren können, die diesen transfiniten Zahlen entsprechen.

Neben der geordneten Summe von Ordnungstypen kann man das *geordnete Produkt* einführen. Es seien M_1 und M_2 Mengen, die Ordnungstypen θ_1 und θ_2 entsprechen. Wir nehmen Exemplare der Menge M_1 (je eines auf jedes Element von M_2) und ersetzen in der Menge M_2 die Elemente durch diese Exemplare von M_1. Die dadurch erhaltene Menge nennt man das geordnete Produkt von M_1 und M_2 und bezeichnet es mit dem Symbol $M_1 \cdot M_2$. Formal wird $M_1 \cdot M_2$ als die Menge der Paare (a, b) mit $a \in M_1$ und $b \in M_2$ konstruiert, wobei $(a_1, b_1) < (a_2, b_2)$ ist, wenn $b_1 < b_2$ (mit beliebigen a_1, a_2) gilt, und $(a_1, b) < (a_2, b)$, wenn $a_1 < a_2$ ist. Analog definiert man das geordnete Produkt einer beliebigen endlichen Anzahl von Faktoren $M_1 \cdot M_2 \cdots M_p$. Den Ordnungstyp des Produktes $M_1 \cdot M_2$ der geordneten Mengen bezeichnet man als *Produkt der Ordnungstypen* θ_1 *und* θ_2:

$$\theta = \theta_1 \cdot \theta_2 .$$

Wie die geordnete Summe, so ist auch das geordnete Produkt nicht kommutativ.

Lemma 2. *Das geordnete Produkt zweier wohlgeordneter Mengen ist eine wohlgeordnete Menge.*

Beweis. Es sei M eine Teilmenge des Produktes $M_1 \cdot M_2$; die Menge M ist eine Menge von Paaren (a, b). Wir betrachten alle zweiten Elemente b der Paare, die zu M gehören. Sie bilden eine gewisse Teilmenge von M_2. Auf Grund der Wohlordnung von M_2 besitzt diese Teilmenge ein erstes Element. Wir bezeichnen es mit b_0 und betrachten alle Paare der Gestalt (a, b_0), die zu M gehören. Ihre ersten Elemente a bilden eine gewisse Teilmenge in M_1. Auf Grund der Wohlordnung von M_1 gibt es unter ihnen ein erstes Element. Wir nennen es a_0. Dann ist das Paar (a_0, b_0), wie man leicht sieht, auch das erste Element von M.

Folgerung. *Das geordnete Produkt von Ordnungszahlen ist eine Ordnungszahl.*

Beispiele. Man sieht leicht, daß $\omega + \omega = \omega \cdot 2$, $\omega + \omega + \omega = \omega \cdot 3$ ist. Es lassen sich auch leicht Mengen konstruieren, die nach den Typen $\omega \cdot n, \omega^2, \omega^2 \cdot n, \omega^3, \ldots, \omega^p, \ldots$ geordnet sind. Alle diese Mengen haben eine abzählbare Mächtigkeit.

Man kann auch andere Operationen mit Ordnungstypen definieren, beispielsweise das Potenzieren, und solche Ordnungstypen wie ω^ω, ω^{ω^ω} usw. betrachten.

1.4.6. Der Vergleich von Ordnungszahlen. Sind n_1 und n_2 zwei endliche Ordnungszahlen, so sind sie entweder gleich, oder die eine ist größer als die andere. Wir wollen jetzt untersuchen, wie es mit der Ordnung der transfiniten Ordnungszahlen aussieht.

Hierzu führen wir folgende Begriffe ein. Jedes Element a einer (linear) geordneten Menge bestimmt einen *Anfangsabschnitt* P (die Gesamtheit der Elemente $< a$) und einen *Rest* Q (die Gesamtheit der Elemente $\geqq a$).

Es seien α und β zwei Ordnungszahlen sowie M und N zwei Mengen vom Typ α bzw. β. Wir setzen $\alpha = \beta$, falls M und N isomorph sind, $\alpha < \beta$, falls M isomorph zu einem passenden Anfangsabschnitt von N ist, und $\alpha > \beta$, falls N isomorph zu einem passenden Anfangsabschnitt von M ist.

Satz 1. *Je zwei Ordnungszahlen α und β sind durch genau eine der drei Beziehungen*

$$\alpha = \beta, \quad \alpha < \beta \quad \textit{und} \quad \alpha > \beta$$

miteinander verknüpft.

Bevor wir mit dem Beweis beginnen, leiten wir folgendes Lemma ab.

Lemma 3. *Ist f eine isomorphe Abbildung einer wohlgeordneten Menge A auf eine Teilmenge B von A, so ist $f(a) \geqq a$ für alle $a \in A$.*

Wir betrachten die Menge der Elemente $a \in A$ mit $f(a) < a$. Wegen der Eigenschaft der Wohlordnung gibt es ein kleinstes Element, das mit a_0 bezeichnet wird. Es sei $b_0 = f(a_0)$. Dann ist $b_0 < a_0$. Da f ein Isomorphismus ist, muß $f(b_0) < f(a_0) = b_0$ sein. Also war a_0 nicht das kleinste Element mit der genannten Eigenschaft.

Aus diesem Lemma folgt sofort, daß eine wohlgeordnete Menge nicht isomorph zu einem ihrer Anfangsabschnitte sein kann. Wäre nämlich A isomorph zu dem Anfangsabschnitt, der durch das Element a bestimmt wird, so müßte $f(a) < a$ sein. Folglich können

$$\alpha = \beta \quad \text{und} \quad \alpha < \beta$$

nicht gleichzeitig richtig sein. Analog sieht man, daß sich $\alpha = \beta$ und $\alpha > \beta$ einander ausschließen, wie auch

$$\alpha < \beta \quad \text{und} \quad \alpha > \beta$$

(anderenfalls würde aus der Transitivität $\alpha < \alpha$ folgen, was nicht sein kann, wie wir eben gesehen haben). Somit haben wir gezeigt, daß das Bestehen einer der Relationen $\alpha \lesseqgtr \beta$ die beiden anderen ausschließt. Wir zeigen jetzt, daß *stets eine dieser Relationen richtig ist,* d. h., daß je zwei Ordnungszahlen miteinander vergleichbar sind.

Hierzu konstruieren wir zu jeder Ordnungszahl α eine Menge $W(\alpha)$, die als ihre „Standarddarstellung" dient. $W(\alpha)$ sei die Menge aller Ordnungszahlen, die kleiner

als α sind. Die Zahlen, die zu $W(\alpha)$ gehören, sind miteinander vergleichbar, und die Menge $W(\alpha)$ (geordnet nach der Größe der Ordnungszahlen) ist vom Typ α. Um diese Aussagen zu beweisen, betrachten wir eine Menge

$$A = \{\ldots, a, \ldots, b, \ldots\}$$

vom Typ α. Nach Definition entsprechen den Ordnungszahlen, die kleiner als α sind, Anfangsabschnitte der Menge A in eineindeutiger Weise. Somit entsprechen diese Ordnungszahlen in eineindeutiger Weise den Elementen der Menge A. Mit anderen Worten, man kann die Elemente einer Menge vom Typ α mit Hilfe der Ordnungszahlen numerieren, die kleiner als α sind,

$$A = (a_0, a_1, \ldots, a_\lambda, \ldots).$$

Es seien jetzt α und β zwei Ordnungszahlen sowie $A = W(\alpha)$ und $B = W(\beta)$ die entsprechenden Mengen vom Typ α und Typ β. Es sei $C = A \cap B$ der Durchschnitt der Mengen A und B, also die Gesamtheit der Ordnungszahlen, die sowohl kleiner als α als auch kleiner als β sind. Die Menge C ist wohlgeordnet, ihren Typ bezeichnen wir mit γ. Wir zeigen, daß $\gamma \leqq \alpha$ gilt. Ist $C = A$, so gilt $\gamma = \alpha$. Ist $C \neq A$, so ist C ein Anfangsabschnitt von A. Um das letzte zu beweisen, betrachten wir zwei beliebige Zahlen $\xi \in C$ und $\eta \in A \setminus C$. Diese Zahlen sind vergleichbar, d. h. $\xi \gtrless \eta$. Die Beziehung $\eta < \xi < \alpha$ ist aber nicht möglich, da sonst η zu C gehören würde. Also ist stets $\xi < \eta$. Hieraus folgt aber, daß C ein Anfangsabschnitt von A ist. Ferner ist γ das erste Element von $A \setminus C$ und somit $\gamma \leqq \alpha$. Entsprechend beweist man $\gamma \leqq \beta$.

Nun ist der Fall $\gamma < \alpha$, $\gamma < \beta$ unmöglich. Denn sonst würde einerseits $\gamma \notin C$ und andererseits

$$\gamma \in A \setminus C, \quad \gamma \in B \setminus C,$$

also $\gamma \in A \cap B = C$ gelten. Folglich sind nur die Fälle

$$\gamma = \alpha,\ \gamma = \beta, \quad \text{also } \alpha = \beta,$$

$$\gamma = \alpha,\ \gamma < \beta, \quad \text{also } \alpha < \beta,$$

$$\gamma < \alpha,\ \gamma = \beta, \quad \text{also } \alpha > \beta$$

denkbar. Somit sind α und β vergleichbar. Damit ist der Satz bewiesen.

Jeder Ordnungszahl entspricht eine bestimmte Mächtigkeit. Aus einem Vergleich der Ordnungszahlen folgt somit ein Vergleich der Mächtigkeiten. Also gilt:

Sind A und B zwei wohlgeordnete Mengen, so sind sie entweder äquivalent (gleichmächtig), oder die Mächtigkeit der einen Menge ist größer als die Mächtigkeit der anderen Menge (d. h., daß wohlgeordnete Mengen niemals unvergleichbare Mächtigkeiten besitzen können).

Wir betrachten die Gesamtheit der Ordnungszahlen, die einer endlichen oder abzählbaren Mächtigkeit entsprechen. Sie bildet eine wohlgeordnete Menge. Man

sieht leicht, daß *diese Menge überabzählbar ist.* Um das zu beweisen, bezeichnen wir in Übereinstimmung mit der üblichen Symbolik den Ordnungstyp der Menge der Ordnungszahlen von Mengen mit abzählbarer Mächtigkeit mit ω_1. Erhält man auf diese Weise eine Menge mit abzählbarer Mächtigkeit, so ist auch eine Menge vom Ordnungstyp $\omega + 1$ abzählbar. Andererseits ist aber klar, daß jede Ordnungszahl, die einer endlichen oder abzählbaren Mächtigkeit entspricht, der Zahl ω_1 vorangeht. Das ist ein Widerspruch. Somit ist ω_1 überabzählbar.

Wir bezeichnen die Mächtigkeit, die der Ordnungszahl ω_1 entspricht, mit $\aleph_1$. Man sieht leicht, daß es keine Mächtigkeit m mit

$$\aleph_0 < m < \aleph_1$$

gibt. Hierzu nehmen wir an, daß m eine derartige Mächtigkeit sei. Dann gibt es in der Menge $W(\omega_1)$ aller Ordnungszahlen, die ω_1 vorangehen, eine wohlgeordnete Untermenge der Mächtigkeit m, die nicht abzählbar ist. Ihr Ordnungstyp α geht der Zahl ω_1 voran; andererseits folgt er auf alle Ordnungszahlen, die zu abzählbaren Mengen gehören. Das ist ein Widerspruch zur Definition von ω_1.

1.4.7. Das Auswahlaxiom, der Satz von Zermelo und andere äquivalente Aussagen. Der Vergleich der Mächtigkeit wohlgeordneter Mengen gibt Anlaß zu der Frage: Kann man jede Menge in irgendeiner Weise wohlordnen? Eine positive Antwort würde unter anderem zur Folge haben, daß es keine unvergleichbaren Mächtigkeiten gibt. Eine solche Antwort stammt von Zermelo, der bewiesen hat, daß *man jede Menge wohlordnen kann.* Der Beweis dieses Satzes (den wir hier nicht bringen werden, man vgl. etwa [2]) verwendet wesentlich das sogenannte *Auswahlaxiom*, das in folgendem besteht.

Es sei A eine Indexmenge, und jedem Index $\alpha \in A$ sei eine Menge M_α zugeordnet. Dann besagt das Auswahlaxiom, daß *man auf A eine Funktion φ konstruieren kann, so daß jedem $\alpha \in A$ ein Element m_α aus M_α entspricht.* Mit anderen Worten, man bildet eine Menge, indem man aus jedem M_α genau ein Element auswählt.

Die Mengenlehre in der hier beschriebenen Form geht auf Cantor und Zermelo zurück und ist eine „naive" Mengenlehre. Das Auswahlaxiom, das man auch das Axiom von Zermelo nennt, entstand im Rahmen der naiven Mengenlehre, wie auch die Kontinuumshypothese. Diese beschäftigt sich mit der Frage, ob die Mächtigkeit des Kontinuums mit der ersten überabzählbaren Mächtigkeit $\aleph_1$ übereinstimmt. Derartige Probleme führten zu zahlreichen Diskussionen und zu einer großen Zahl von Arbeiten über mathematische Logik und über die Grundlagen der Mathematik. Durch Gödel-Bernays und Zermelo-Fraenkel entstanden axiomatische Mengenlehren. Wir verweisen den Leser auf die einschlägige Literatur[1]). Wir bemerken noch, daß ein Verzicht auf das Auswahlaxiom zu einer wesentlichen Verarmung der Mengenlehre führt. Die Kritik an der naiven Mengenlehre und die Versuche, das Auswahlaxiom zu vermeiden, haben zu bemerkenswerten Theorien geführt, wie der Theorie der rekursiven Funktionen und dem Begriff der berechenbaren Zahl.

---

[1]) Vgl. [16] sowie P. Cohen, Set Theory and the Continuum Hypothesis, W. A. Benjamin, Inc., New York—Amsterdam. Anm. d. Übers.: Man vergleiche auch [33].

Wir formulieren einige Aussagen, von denen jede äquivalent zum Auswahlaxiom ist (d. h., daß jede von ihnen mit Hilfe des Auswahlaxioms bewiesen werden kann und daß man umgekehrt das Auswahlaxiom beweisen kann, wenn man irgendeine von diesen Aussagen als richtig voraussetzt). Wir bemerken, daß der Satz von ZERMELO eine derartige Aussage ist. Setzt man voraus, daß die Mengen M_α wohlgeordnet sind, so kann man die Funktion φ aus dem Auswahlaxiom konstruieren, indem man aus jeder Menge M_α das erste Element nimmt.

Zur Formulierung anderer Aussagen, die äquivalent zum Auswahlaxiom sind, führen wir folgende Begriffe ein. Es sei M eine halbgeordnete Menge. Eine Untermenge A von M wird als *Kette* bezeichnet, wenn je zwei Elemente aus A vergleichbar sind (im Sinne der Halbordnung auf M). Eine Kette heißt *maximal*, wenn sie nicht als echte Untermenge in einer anderen Kette aus M enthalten ist. Ein Element α aus einer halbgeordneten Menge M heißt *obere Schranke* der Untermenge $M' \subset M$, falls jedes Element $a' \in M'$ unterhalb von α liegt.

Satz von HAUSDORFF. *In einer halbgeordneten Menge ist jede Kette in einer maximalen Kette enthalten.*

Die folgende Aussage ist die nützlichste unter allen äquivalenten Formulierungen des Auswahlaxioms.

Lemma von ZORN. *Falls jede Kette einer halbgeordneten Menge M eine obere Schranke besitzt, liegt jedes Element aus M unterhalb eines maximalen Elementes.*

Einen Beweis der Äquivalenz aller hier angegebenen Aussagen (Auswahlaxiom, Satz von ZERMELO, Satz von HAUSDORFF, Lemma von ZORN) findet man z. B. in dem Buch von А. Г. Курош, Лекции по общей алгебре, Физматгиз, 1962; man vergleiche auch [7]. Wir werden jetzt nicht darauf eingehen.

Wenn die Menge der oberen Schranken der Untermenge A ein kleinstes Element a besitzt, nennt man a die *obere Grenze* (das *Supremum*) der Untermenge A. Analog definiert man die *untere Grenze* (das *Infimum*). Eine halbgeordnete Menge, deren jede *endliche* nichtleere Untermenge eine obere Grenze und eine untere Grenze besitzt, nennt man einen *Verband*.

1.4.8. Transfinite Induktion. Die vollständige Induktion ist eine weit verbreitete Methode zum Beweis von Aussagen. Sie besteht bekanntlich in folgendem. Gegeben sei eine Behauptung $P(n)$, die für jede natürliche Zahl n formuliert wird. Ferner sei bekannt, daß

1. die Behauptung $P(1)$ richtig ist,

2. aus der Richtigkeit von $P(k)$ für alle $k \leqq n$ folgt, daß auch $P(n + 1)$ richtig ist.

Dann ist die Behauptung $P(n)$ für alle $n = 1, 2, \ldots$ richtig. Anderenfalls würde man unter denjenigen n, für die $P(n)$ nicht richtig ist, eine kleinste Zahl finden, die mit n_1 bezeichnet wird. Da $n_1 > 1$ ist und somit $n_1 - 1$ ebenfalls eine natürliche Zahl ist, erhält man einen Widerspruch zu Bedingung 2.

Das gleiche Verfahren kann man benutzen, wenn man die natürlichen Zahlen durch eine beliebige wohlgeordnete Menge ersetzt. Man spricht dann von *transfiniter Induktion*. Die Methode der transfiniten Induktion besteht somit in folgendem. Vorgegeben ist eine wohlgeordnete Menge A (man kann etwa die Menge der Ordnungszahlen nehmen, die kleiner als eine vorgegebene Ordnungszahl sind). Ferner sei $P(a)$ eine Behauptung, die für alle $a \in A$ formuliert wird. Die Behauptung $P(a)$ sei richtig für das erste Element aus A und für a, sofern sie für alle Elemente richtig ist, die a vorangehen. Dann ist $P(a)$ für alle $a \in A$ richtig. Nimmt man an, daß es Elemente aus A gibt, für die $P(a)$ nicht richtig ist, so gibt es in der Menge dieser Elemente ein kleinstes Element, das mit a^* bezeichnet werde. Da die Behauptung für alle $a < a^*$ richtig ist, kommen wir zu einem Widerspruch.

Da man nach dem Satz von Zermelo jede Menge wohlordnen kann, läßt sich im Prinzip die transfinite Induktion auf jede Menge anwenden. In den Anwendungen ist es aber bequemer, das Lemma von Zorn zu benutzen, da man hierzu nur die Halbordnung der betrachteten Menge benötigt. Bei vielen Aufgaben, die die Anwendung des Lemmas von Zorn erfordern, gibt es in natürlicher Weise eine Halbordnung der betrachteten Objekte.

1.5. Mengensysteme[1])

1.5.1. Mengenringe. Jede Menge, deren Elemente selbst irgendwelche Mengen sind, bezeichnen wir als *Mengensystem*. Zur formalen Kennzeichnung von Mengensystemen verwenden wir große Frakturbuchstaben. Wenn nichts Gegenteiliges vereinbart ist, betrachten wir stets Mengensysteme, deren Elemente Untermengen einer festen Menge X sind. Unser Hauptinteresse gilt dabei Mengensystemen, die gegenüber den in 1.1. definierten Mengenoperationen in gewissem Sinne abgeschlossen sind.

Definition 1. Ein nichtleeres Mengensystem $\mathfrak{R}$ heißt *Mengenring* (*Ring*), wenn mit zwei Mengen A und B stets auch ihre symmetrische Differenz und ihr Durchschnitt zum Mengensystem gehören, d. h., wenn aus $A \in \mathfrak{R}$ und $B \in \mathfrak{R}$ auch $A \triangle B \in \mathfrak{R}$ und $A \cap B \in \mathfrak{R}$ folgt.

Weil für beliebige Mengen A und B

$$A \cup B = (A \triangle B) \triangle (A \cap B)$$

und

$$A \setminus B = A \triangle (A \cap B)$$

[1]) In diesem Abschnitt werden Begriffe betrachtet, die erst in Kapitel 5 bei der Darlegung der allgemeinen Maßtheorie benutzt werden. Daher kann die Lektüre dieses Abschnitts bis 5.2. aufgeschoben werden. Der Leser, der sich nur für Maßtheorie ebener Mengen (5.1.) interessiert, kann diesen Abschnitt ganz überschlagen.

ist, folgt aus der Zugehörigkeit von A und B zum Ring $\mathfrak{R}$ auch $A \cup B \in \mathfrak{R}$ und $A \setminus B \in \mathfrak{R}$. Ein Ring ist also ein Mengensystem, das bezüglich der Bildung von Vereinigung, Durchschnitt, Differenz und symmetrischer Differenz abgeschlossen ist. Es ist klar, daß damit ein Ring auch gegenüber der Bildung einer beliebigen endlichen Anzahl von Vereinigungen und Durchschnitten der Form

$$C = \bigcup_{k=1}^{n} A_k, \qquad D = \bigcap_{k=1}^{n} A_k$$

abgeschlossen ist.

Ein beliebiger Ring enthält stets die leere Menge, weil $A \setminus A = \emptyset$ ist. Das Mengensystem, das nur die leere Menge als Element enthält, ist der kleinstmögliche Mengenring.

Eine Menge E heißt *Eins* des Mengensystems $\mathfrak{S}$, wenn E selbst zu $\mathfrak{S}$ gehört und für jede Menge $A \in \mathfrak{S}$ die Beziehung

$$A \cap E = A$$

erfüllt ist.

Die Eins eines Mengensystems $\mathfrak{S}$ ist nichts anderes als die größte Menge des Systems, die alle anderen Mengen aus $\mathfrak{S}$ enthält.

Ein Mengenring mit Eins heißt *Mengenalgebra* (*Algebra*).

Beispiele

1. Für eine beliebige Menge A ist das System $\mathfrak{M}(A)$ aller Untermengen von A eine Mengenalgebra mit der Eins $E = A$.

2. Für eine beliebige nichtleere Menge A ist das System $\{\emptyset, A\}$, das aus der Menge A und der leeren Menge $\emptyset$ besteht, eine Mengenalgebra mit der Eins $E = A$.

3. Das System aller endlichen Untermengen einer beliebigen Menge A ist ein Mengenring. Dieser Ring ist dann und nur dann eine Algebra, wenn A selbst endlich ist.

4. Das System aller beschränkten Untermengen der Zahlengeraden ist ein Ring ohne Eins.

Aus der Definition des Mengenrings folgt unmittelbar

Satz 1. *Der Durchschnitt* $\mathfrak{R} = \bigcap_{\alpha} \mathfrak{R}_\alpha$ *einer beliebigen Menge von Ringen ist ein Ring.*

Eine einfache, aber für das Weitere wichtige Tatsache ist folgender Satz.

Satz 2. *Für ein beliebiges nichtleeres Mengensystem* $\mathfrak{S}$ *existiert genau ein Ring* $\mathfrak{R}(\mathfrak{S})$, *der* $\mathfrak{S}$ *enthält und seinerseits in jedem* $\mathfrak{S}$ *enthaltenden Ring* $\mathfrak{R}$ *liegt.*

Beweis. Aus der Definition von $\mathfrak{R}(\mathfrak{S})$ folgt sofort, daß $\mathfrak{R}(\mathfrak{S})$ durch das System $\mathfrak{S}$ eindeutig bestimmt wird. Um die Existenz des Ringes $\mathfrak{R}(\mathfrak{S})$ nachzuweisen, betrachten

wir die Vereinigung $X = \bigcap_{A \in \mathfrak{S}} A$ aller Mengen A aus $\mathfrak{S}$ und den Ring $\mathfrak{M}(X)$ aller Untermengen von X. Bezeichnet Σ die Gesamtheit aller Mengenringe, die in $\mathfrak{M}(X)$ enthalten sind und ihrerseits das Mengensystem $\mathfrak{S}$ enthalten, so ist der Durchschnitt

$$\mathfrak{P} = \bigcap_{\mathfrak{R} \in \Sigma} \mathfrak{R}$$

aller dieser Ringe offensichtlich der gesuchte Ring $\mathfrak{R}(\mathfrak{S})$. Für einen beliebigen Ring $\mathfrak{R}^*$, der $\mathfrak{S}$ enthält, ist nämlich der Durchschnitt $\mathfrak{R} = \mathfrak{R}^* \cap \mathfrak{M}(X)$ ein Ring aus Σ und

$$\mathfrak{S} \subset \mathfrak{R} \subset \mathfrak{R}^*.$$

Also erfüllt $\mathfrak{P}$ tatsächlich die Forderung der Minimalität. Dieser Ring $\mathfrak{P}$ heißt deshalb *minimaler Ring über* $\mathfrak{S}$ oder *von* $\mathfrak{S}$ *erzeugter Ring* und wird mit $\mathfrak{R}(\mathfrak{S})$ bezeichnet.

1.5.2. Semiringe von Mengen. In einer Reihe von Fragen, zum Beispiel in der Maßtheorie, spielt neben dem Begriff des Ringes der allgemeinere Begriff des Semiringes von Mengen eine wichtige Rolle.

Definition 2. Ein Mengensystem $\mathfrak{S}$ heißt *Semiring*, wenn es die leere Menge $\emptyset$ enthält, bezüglich der Bildung von Durchschnitten abgeschlossen ist und folgende Eigenschaft besitzt: Für beliebige Mengen A und $A_1 \subset A$ aus $\mathfrak{S}$ gibt es eine Darstellung von A in der Form $A = \bigcup_{k=1}^{n} A_k$, wobei A_k paarweise disjunkte Mengen aus $\mathfrak{S}$ sind und die erste Menge die vorgegebene Untermenge A_1 von A ist.

Im weiteren bezeichnen wir jedes System von paarweise disjunkten Mengen $A_1, A_2, \ldots, A_n$, deren Vereinigung eine vorgegebene Menge A ist, als *endliche Zerlegung* der Menge A.

Jeder Mengenring $\mathfrak{R}$ ist ein Semiring, denn wenn A und $A_1 \subset A$ in $\mathfrak{R}$ liegen, ist

$$A = A_1 \cup A_2 \quad \text{mit} \quad A_2 = A \setminus A_1 \in \mathfrak{R}$$

eine Zerlegung von A, wie sie die Definition des Semiringes fordert.

Ein Beispiel für einen Semiring, der nicht zugleich Ring ist, ist die Gesamtheit aller Intervalle (a, b), Segmente $[a, b]$ und Halbsegmente $[a, b)$ bzw. $(a, b]$ auf der Zahlengeraden[1]), andere Beispiele sind die Gesamtheit aller „halboffenen" Rechtecke $a < x \leqq b$, $c < y \leqq d$ in der Ebene und die Gesamtheit aller halboffenen Parallelepipede im Raum.

Semiringe von Mengen besitzen folgende Eigenschaften.

[1]) In dieser Gesamtheit sind natürlich auch das leere Intervall (a, a) und das Segment $[a, a]$, das nur aus einem Punkt besteht, enthalten.

Lemma 1. *Ist $\mathfrak{S}$ ein Semiring und sind $A_1, A_2, \ldots, A_n \in \mathfrak{S}$ paarweise disjunkte Untermengen einer Menge $A \in \mathfrak{S}$, dann kann man das System der Mengen A_i $(i = 1, 2, \ldots, n)$ durch Mengen $A_{n+1}, \ldots, A_s \in \mathfrak{S}$ zu einer endlichen Zerlegung*

$$A = \bigcup_{k=1}^{s} A_k \qquad (s \geqq n),$$

von A in $\mathfrak{S}$ ergänzen.

Beweis. Wir beweisen die Aussage durch vollständige Induktion. Für $n = 1$ folgt die Richtigkeit der Aussage des Lemmas unmittelbar aus der Definition des Semiringes. Setzen wir nun voraus, daß die Aussage für $n \leqq m$ richtig ist und $A_1, \ldots, A_m, A_{m+1}$ Mengen aus $\mathfrak{S}$ sind, die die Voraussetzungen des Lemmas erfüllen, dann existiert eine solche Zerlegung von A,

$$A = A_1 \cup A_2 \cup \cdots \cup A_m \cup B_1 \cup B_2 \cup \cdots \cup B_p,$$

für die $B_q \in \mathfrak{S}$ $(q = 1, 2, \ldots, p)$ ist. Setzen wir

$$B_{q1} = A_{m+1} \cap B_q,$$

dann gibt es nach Definition des Semiringes $\mathfrak{S}$ eines solche Zerlegung von B_q,

$$B_q = B_{q1} \cup B_{q2} \cup \cdots \cup B_{qr_q},$$

für die alle B_{qj} zu $\mathfrak{S}$ gehören. Es ist nun offensichtlich

$$A = A_1 \cup \cdots \cup A_m \cup A_{m+1} \cup \bigcup_{q=1}^{p} \left(\bigcup_{j=2}^{r_q} B_{qj} \right).$$

Die Existenz dieser Mengen $B_{qj} \in \mathfrak{S}$, die paarweise disjunkt sind und die Mengen $A_1, \ldots, A_m, A_{m+1}$ zu einer endlichen Zerlegung von A in $\mathfrak{S}$ ergänzen, ergibt die Aussage des Lemmas für $n = m + 1$. Damit ist diese Aussage nach dem Induktionsprinzip für alle natürlichen Zahlen n richtig.

Lemma 2. *Ist $\mathfrak{S}$ ein Semiring und $A_1, \ldots, A_n$ ein beliebiges endliches System von Mengen aus $\mathfrak{S}$, dann gibt es in $\mathfrak{S}$ ein endliches System von paarweise disjunkten Mengen $B_1, \ldots, B_t$, so daß jede Menge A_k als Vereinigung von gewissen Mengen B_s,*

$$A_k = \bigcup_{s \in M_k} B_s,$$

dargestellt werden kann.

Beweis. Für $n = 1$ ist das Lemma trivial, weil einfach $t = 1$, $B_1 = A_1$ gesetzt werden kann. Setzen wir nun voraus, daß das Lemma für $n \leqq m$ richtig und $A_1, \ldots, A_m, A_{m+1}$ ein beliebiges System von Mengen aus $\mathfrak{S}$ ist. Es seien $B_1, \ldots, B_t$ diejenigen Mengen aus $\mathfrak{S}$, die die Aussage des Lemmas bezüglich $A_1, \ldots, A_m$ erfüllen, und

$$B_{s1} = A_{m+1} \cap B_s \qquad (s = 1, \ldots, t).$$

Dann gibt es nach Lemma 1 für die Menge A_{m+1} eine Zerlegung in $\mathfrak{S}$, die die Mengen B_{s1} $(s = 1, \ldots, t)$ enthält,

$$A_{m+1} = \bigcup_{s=1}^{t} B_{s1} \cup \bigcup_{p=1}^{q} B_p', \quad B_p' \in \mathfrak{S}, \tag{1}$$

und nach Definition des Semiringes für die Mengen B_s $(s = 1, \ldots, t)$ jeweils auch eine Zerlegung in $\mathfrak{S}$, die die Menge B_{s1} enthält,

$$B_s = B_{s1} \cup B_{s2} \cup \cdots \cup B_{sr_s}, \quad B_{sj} \in \mathfrak{S}.$$

Man prüft leicht nach, daß

$$A_k = \bigcup_{s \in M_k} \bigcup_{j=1}^{r_s} B_{sj} \quad (k = 1, 2, \ldots, m)$$

und daß alle Mengen

$$B_{sj}, \; B_p'$$

paarweise disjunkt sind. Daraus folgt zusammen mit (1), daß die Mengen $B_{sj}, B_p' \in \mathfrak{S}$ ein Mengensystem bilden, das die Aussage des Lemmas bezüglich $A_1, \ldots, A_m, A_{m+1}$ erfüllt. Damit ist das Lemma nach dem Induktionsprinzip für alle n bewiesen.

1.5.3. Ringe, die von Semiringen erzeugt werden. Wie wir schon in 1.5.1. festgestellt haben, existiert für jedes Mengensystem $\mathfrak{S}$ ein eindeutig bestimmter minimaler Ring, der $\mathfrak{S}$ enthält. Die explizite Konstruktion dieses Ringes $\mathfrak{R}(\mathfrak{S})$ ist für ein beliebiges Mengensystem $\mathfrak{S}$ jedoch ziemlich schwierig. Vollständig überschaubar ist sie in dem wichtigen Fall, in dem $\mathfrak{S}$ ein Semiring ist. Der folgende Satz beschreibt diese Konstruktion.

Satz 3. *Wenn $\mathfrak{S}$ ein Semiring ist, stimmt $\mathfrak{R}(\mathfrak{S})$ mit dem System $\mathfrak{Z}$ derjenigen Mengen A überein, die eine endliche Zerlegung in $\mathfrak{S}$,*

$$A = \bigcup_{k=1}^{n} A_k, \quad A_k \in \mathfrak{S},$$

gestatten.

Beweis. Wir zeigen zuerst, daß das System $\mathfrak{Z}$ ein Ring ist. Wenn A und B zwei beliebige Mengen aus $\mathfrak{Z}$ sind, dann besitzen sie Zerlegungen in $\mathfrak{S}$,

$$A = \bigcup_{i=1}^{n} A_i, \quad B = \bigcup_{j=1}^{m} B_j, \quad A_i \in \mathfrak{S}, \quad B_j \in \mathfrak{S}.$$

Weil $\mathfrak{S}$ ein Semiring ist, gehören die Mengen

$$C_{ij} = A_i \cap B_j$$

auch zu $\mathfrak{S}$. Nach Lemma 1 gibt es Zerlegungen von A_i und B_j in $\mathfrak{S}$ der Form

$$\begin{aligned} A_i &= \bigcup_j C_{ij} \cup \bigcup_{k=1}^{r_i} D_{ik}, \quad D_{ik} \in \mathfrak{S}; \\ B_j &= \bigcup_i C_{ij} \cup \bigcup_{l=1}^{s_j} E_{jl}, \quad E_{jl} \in \mathfrak{S}. \end{aligned} \tag{2}$$

Damit ergibt sich für die Mengen $A \cap B$ und $A \triangle B$

$$A \cap B = \bigcup_{i,j} C_{ij},$$

$$A \triangle B = \bigcup_{i,k} D_{ik} \cup \bigcup_{j,l} E_{jl}$$

eine Darstellung durch endlich viele paarweise disjunkte Mengen C_{ij} bzw. D_{ik} und E_{jl} aus $\mathfrak{S}$. Das bedeutet, daß $A \cap B$ und $A \triangle B$ zu $\mathfrak{Z}$ gehören, $\mathfrak{Z}$ also ein Ring ist. Da der Ring $\mathfrak{Z}$ unter allen das System $\mathfrak{S}$ enthaltenden Ringen offensichtlich der kleinste ist, ergibt sich $\mathfrak{Z} = \mathfrak{R}(\mathfrak{S})$, was zu zeigen war.

1.5.4. σ-Algebren. Bei verschiedenen Fragestellungen, insbesondere in der Maßtheorie, ist es notwendig, Vereinigungen und Durchschnitte nicht nur für endlich viele Mengen, sondern auch für abzählbar viele Mengen zu betrachten. Deshalb ist es zweckmäßig, neben dem Begriff des Mengenringes noch folgende Begriffe einzuführen.

Definition 3. Ein Mengenring heißt *σ-Ring*, wenn er mit jeder Folge von Mengen $A_1, A_2, \ldots, A_n, \ldots$ auch stets deren Vereinigung

$$S = \bigcup_n A_n$$

enthält.

Definition 4. Ein Mengenring heißt *δ-Ring*, wenn er mit jeder Folge von Mengen $A_1, A_2, \ldots, A_n, \ldots$ auch stets deren Durchschnitt

$$D = \bigcap_n A_n$$

enthält.

Entsprechend heißt ein σ-Ring mit Eins *σ-Algebra* und ein δ-Ring mit Eins *δ-Algebra*. Aus den Dualitätsbeziehungen

$$\bigcup_n A_n = E \setminus \bigcap_n (E \setminus A_n),$$

$$\bigcap_n A_n = E \setminus \bigcup_n (E \setminus A_n)$$

(vgl. 1.1.) folgt sofort, daß diese beiden Begriffe übereinstimmen, d. h., jede σ-Algebra ist gleichzeitig eine δ-Algebra, und jede δ-Algebra ist eine σ-Algebra.

Das einfachste Beispiel für eine σ-Algebra ist die Gesamtheit aller Untermengen einer beliebigen Menge A.

Für ein beliebiges Mengensystem $\mathfrak{S}$ existiert immer mindestens eine σ-Algebra, die dieses System enthält. Setzen wir nämlich

$$X = \bigcup_{A \in \mathfrak{S}} A,$$

so ist die Gesamtheit $\mathfrak{B}$ aller Untermengen von X eine σ-Algebra, die das System $\mathfrak{S}$ enthält.

Ist $\tilde{\mathfrak{B}}$ eine beliebige σ-Algebra, die das Mengensystem $\mathfrak{S}$ enthält, und $\tilde{X}$ ihre Eins, dann ist jede Menge $A \in \mathfrak{S}$ in $\tilde{X}$ enthalten und folglich auch $X = \bigcup_{A \in \mathfrak{S}} A \subset \tilde{X}$. Wir nennen die σ-Algebra $\tilde{\mathfrak{B}}$ *irreduzibel* bezüglich des Systems $\mathfrak{S}$, wenn $\tilde{X} = \bigcup_{A \in \mathfrak{S}} A$, d. h. die irreduzible σ-Algebra ist diejenige $\mathfrak{S}$ umfassende σ-Algebra, deren Mengen keine anderen Elemente enthalten, als in den Mengen des Systems $\mathfrak{S}$ enthalten sind. Es ist klar, daß wir unsere Betrachtungen in jedem Fall auf solche σ-Algebren einschränken können.

Für irreduzible σ-Algebren ist ein Satz, analog dem Satz 2 für Ringe, richtig.

Satz 4. *Für ein beliebiges nichtleeres Mengensystem $\mathfrak{S}$ existiert eine irreduzible σ-Algebra $\mathfrak{B}(\mathfrak{S})$, die $\mathfrak{S}$ enthält und ihrerseits in jeder $\mathfrak{S}$ enthaltenden σ-Algebra liegt.*

Beweis. Satz 4 wird völlig analog zu Satz 2 bewiesen. Die σ-Algebra $\mathfrak{B}(\mathfrak{S})$ heißt *minimale σ-Algebra über $\mathfrak{S}$*.

In der Analysis spielen die sogenannten *Borelmengen* (oder *B-Mengen*) der Zahlengeraden eine wichtige Rolle, sie gehören zur minimalen σ-Algebra über der Gesamtheit aller Segmente $[a, b]$.

1.5.5. Mengensysteme und Abbildungen. Es sei $y = f(x)$ eine Funktion, die auf der Menge M definiert ist und deren Funktionswerte die Menge N bilden. Für ein beliebiges System $\mathfrak{M}$ von Untermengen der Menge M bezeichnen wir mit $f(\mathfrak{M})$ das Mengensystem aller Bilder von Mengen aus $\mathfrak{M}$ und für ein beliebiges System $\mathfrak{N}$ von Untermengen der Menge N mit $f^{-1}(\mathfrak{N})$ das Mengensystem aller Urbilder von Mengen aus $\mathfrak{N}$. Mit diesen Bezeichnungen sind nun folgende Aussagen richtig, die bei der späteren Untersuchung von meßbaren Funktionen (in 5.4.) benutzt werden:

1. Wenn $\mathfrak{N}$ ein Ring ist, dann ist auch $f^{-1}(\mathfrak{N})$ ein Ring.
2. Wenn $\mathfrak{N}$ eine Algebra ist, dann ist auch $f^{-1}(\mathfrak{N})$ eine Algebra.
3. Wenn $\mathfrak{N}$ eine σ-Algebra ist, dann ist auch $f^{-1}(\mathfrak{N})$ eine σ-Algebra.
4. $\mathfrak{R}(f^{-1}(\mathfrak{N})) = f^{-1}(\mathfrak{R}(\mathfrak{N}))$.
5. $\mathfrak{B}(f^{-1}(\mathfrak{N})) = f^{-1}(\mathfrak{B}(\mathfrak{N}))$.

Den Beweis für diese Aussagen überlassen wir dem Leser, der auch prüfen sollte, ob die Aussagen richtig bleiben, wenn f^{-1} durch f und $\mathfrak{N}$ durch $\mathfrak{M}$ ersetzt werden.

2. Metrische und topologische Räume

2.1. Der Begriff des metrischen Raumes

2.1.1. Definition und grundlegende Beispiele. Der Grenzübergang ist eine der wichtigsten Operationen der Analysis. Dieser Operation liegt die Tatsache zugrunde, daß zwischen zwei Punkten auf der Zahlengeraden ein Abstand erklärt ist. Viele fundamentale Sachverhalte der Analysis sind nicht mit der algebraischen Natur der reellen Zahlen verbunden (d. h. damit, daß sie einen Körper bilden), sondern stützen sich nur auf den Begriff des Abstandes. Durch Verallgemeinerung der Betrachtungsweise der reellen Zahlen als einer Menge, in der ein Abstand zwischen den Elementen erklärt ist, gelangen wir zum Begriff des metrischen Raumes, einem der wichtigsten Begriffe der modernen Mathematik. Im folgenden werden wir die Grundzüge der Theorie der metrischen Räume und ihrer Verallgemeinerung, der topologischen Räume, darlegen. Die Resultate dieses Kapitels sind grundlegend für die gesamten weiteren Ausführungen.

Definition. Als *metrischen Raum* bezeichnen wir ein Paar (X, ϱ), das aus einer *Menge* (einem Raum) X von *Elementen* (Punkten) und einem *Abstand*, einer *Metrik*, besteht, d. h. einer eindeutigen nichtnegativen reellen Funktion $\varrho(x, y)$, die für beliebige x und y aus X erklärt ist und folgenden drei Axiomen genügt:

1. $\varrho(x, y) = 0$ gilt dann und nur dann, wenn $x = y$ ist,
2. $\varrho(x, y) = \varrho(y, x)$ (Symmetrie),
3. $\varrho(x, z) \leqq \varrho(x, y) + \varrho(y, z)$ (Dreiecksungleichung).

Einen metrischen Raum, d. h. ein Paar (X, ϱ), werden wir gewöhnlich mit einem einzigen Buchstaben bezeichnen:

$$R = (X, \varrho).$$

Sind Mißverständnisse ausgeschlossen, so bezeichnen wir einen metrischen Raum oft auch mit demselben Symbol wie seine „Grundmenge" X.

Wir führen nun Beispiele metrischer Räume an. Einige von diesen Räumen spielen in der Analysis eine sehr wichtige Rolle.

1. Setzen wir für die Elemente einer beliebigen Menge

$$\varrho(x, y) = \begin{cases} 0 \text{ für } x = y, \\ 1 \text{ für } x \neq y, \end{cases}$$

so erhalten wir offenbar einen metrischen Raum. Einen derartigen Raum nennt man *Raum isolierter Punkte.*

2. Die Menge der reellen Zahlen mit dem Abstand

$$\varrho(x, y) = |x - y|$$

bildet den metrischen Raum $\boldsymbol{R}^1$.

3. Die Menge aller n-Tupel von reellen Zahlen

$$x = (x_1, x_2, \ldots, x_n)$$

mit dem Abstand

$$\varrho(x, y) = \sqrt{\sum_{k=1}^{n} (y_k - x_k)^2} \tag{1}$$

heißt *n-dimensionaler euklidischer Raum* $\boldsymbol{R}^n$. Die Gültigkeit der Axiome 1 und 2 für den $\boldsymbol{R}^n$ ist trivial. Wir zeigen, daß im $\boldsymbol{R}^n$ auch die Dreiecksungleichung gilt.

Es sei $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ und $z = (z_1, \ldots, z_n)$; dann lautet die Dreiecksungleichung

$$\sqrt{\sum_{k=1}^{n} (z_k - x_k)^2} \leqq \sqrt{\sum_{k=1}^{n} (y_k - x_k)^2} + \sqrt{\sum_{k=1}^{n} (z_k - y_k)^2}\,. \tag{2}$$

Wenn wir $y_k - x_k = a_k$, $z_k - y_k = b_k$ setzen, ist $z_k - x_k = a_k + b_k$, und die Ungleichung (2) geht über in

$$\sqrt{\sum_{k=1}^{n} (a_k + b_k)^2} \leqq \sqrt{\sum_{k=1}^{n} a_k^2} + \sqrt{\sum_{k=1}^{n} b_k^2}\,. \tag{3}$$

Diese Ungleichung folgt aber sofort aus der bekannten Schwarzschen (Cauchy-Bunjakovskijschen) Ungleichung[1])

$$\left(\sum_{k=1}^{n} a_k b_k\right)^2 \leqq \sum_{k=1}^{n} a_k^2 \cdot \sum_{k=1}^{n} b_k^2\,. \tag{4}$$

[1]) Die Schwarzsche Ungleichung folgt aus der Identität

$$\left(\sum_{k=1}^{n} a_k b_k\right)^2 = \sum_{k=1}^{n} a_k^2 \sum_{k=1}^{n} b_k^2 - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} (a_i b_j - b_i a_j)^2,$$

die man unmittelbar nachprüft.

Mit der Schwarzschen Ungleichung ergibt sich

$$\begin{aligned}\sum_{k=1}^{n}(a_k+b_k)^2 &= \sum_{k=1}^{n}a_k^2+2\sum_{k=1}^{n}a_kb_k+\sum_{k=1}^{n}b_k^2\\ &\leqq \sum_{k=1}^{n}a_k^2+2\sqrt{\sum_{k=1}^{n}a_k^2\cdot\sum_{k=1}^{n}b_k^2}+\sum_{k=1}^{n}b_k^2\\ &= \left(\sqrt{\sum_{k=1}^{n}a_k^2}+\sqrt{\sum_{k=1}^{n}b_k^2}\right)^2.\end{aligned}$$

Damit ist (3) und folglich auch (2) gezeigt.

4. Wir betrachten dieselbe Menge aller n-Tupel $x = (x_1, \ldots, x_n)$ von reellen Zahlen, aber definieren den Abstand durch die Formel

$$\varrho_1(x, y) = \sum_{k=1}^{n} |x_k - y_k|. \tag{5}$$

Die Gültigkeit der Axiome 1 bis 3 ist hier offensichtlich. Wir bezeichnen diesen metrischen Raum mit dem Symbol $\boldsymbol{R}_1^n$.

5. Wir nehmen wiederum dieselbe Menge wie schon in den Beispielen 3 und 4 und definieren den Abstand zwischen ihren Elementen durch die Formel

$$\varrho_0(x, y) = \max_{1\leqq k\leqq n} |y_k - x_k|. \tag{6}$$

Die Gültigkeit der Axiome 1 bis 3 ist wieder trivial. Dieser Raum, den wir mit $\boldsymbol{R}_0^n$ bezeichnen, ist in vielen Fragen der Analysis genauso gut geeignet wie der euklidische Raum $\boldsymbol{R}^n$.

Die letzten drei Beispiele zeigen, daß es manchmal wichtig ist, für den metrischen Raum und für die Menge seiner Punkte verschiedene Bezeichnungen zu haben, weil ein und dieselbe Trägermenge verschieden metrisiert werden kann.

6. Die Menge $C[a, b]$ aller stetigen reellwertigen Funktionen, die auf dem Intervall $[a, b]$ definiert sind, bilden mit dem Abstand

$$\varrho(f, g) = \max_{a\leqq t\leqq b} |g(t) - f(t)| \tag{7}$$

ebenfalls einen metrischen Raum. Die Axiome 1 bis 3 beweist man unmittelbar. Dieser Raum spielt eine sehr wichtige Rolle in der Analysis. Wir werden ihn mit demselben Symbol $C[a, b]$ bezeichnen wie die Menge der Punkte dieses Raumes. Anstelle von $C[0, 1]$ werden wir einfach C schreiben.

7. Mit l_2 bezeichnen wir den metrischen Raum, dessen Punkte alle Folgen

$$x = (x_1, x_2, \ldots, x_n, \ldots)$$

von reellen Zahlen sind, die die Bedingung

$$\sum_{k=1}^{\infty} x_k^2 < \infty$$

erfüllen, und definieren den Abstand durch die Formel

$$\varrho(x, y) = \sqrt{\sum_{k=1}^{\infty} (y_k - x_k)^2}. \tag{8}$$

Aus der elementaren Ungleichung

$$(x_k \pm y_k)^2 \leqq 2(x_k^2 + y_k^2)$$

folgt, daß die Funktion $\varrho(x, y)$ für beliebige $x, y \in l_2$ einen Sinn hat, d. h., daß die Reihe $\sum_{k=1}^{\infty} (y_k - x_k)^2$ konvergiert, wenn

$$\sum_{k=1}^{\infty} x_k^2 < \infty \quad \text{und} \quad \sum_{k=1}^{\infty} y_k^2 < \infty$$

ist. Wir zeigen jetzt, daß die Funktion (8) die Axiome eines metrischen Raumes erfüllt. Die Axiome 1 und 2 gelten offensichtlich, und die Dreiecksungleichung nimmt hier die Form

$$\sqrt{\sum_{k=1}^{\infty} (z_k - x_k)^2} \leqq \sqrt{\sum_{k=1}^{\infty} (z_k - y_k)^2} + \sqrt{\sum_{k=1}^{\infty} (y_k - x_k)^2} \tag{9}$$

an. Jede der drei hier auftretenden Reihen konvergiert nach dem oben Gesagten. Andererseits ist für jedes n die Ungleichung

$$\sqrt{\sum_{k=1}^{n} (z_k - x_k)^2} \leqq \sqrt{\sum_{k=1}^{n} (z_k - y_k)^2} + \sqrt{\sum_{k=1}^{n} (y_k - x_k)^2}$$

gültig (vgl. Beispiel 4). Durch den Grenzübergang $n \to \infty$ erhalten wir daraus (9), d. h. die Dreiecksungleichung in l_2.

8. Wir betrachten wie in Beispiel 6 die Gesamtheit aller auf dem Intervall $[a, b]$ stetigen Funktionen. Jedoch definieren wir den Abstand anders, und zwar setzen wir

$$\varrho(x, y) = \left(\int_a^b \big(x(t) - y(t)\big)^2 \, dt\right)^{1/2}. \tag{10}$$

Diesen metrischen Raum werden wir mit $C^2[a, b]$ bezeichnen und ihn *Raum der stetigen Funktionen mit der quadratischen Metrik* nennen. Hier gelten offenbar wieder die Axiome 1 und 2 eines metrischen Raumes, und die Dreiecksungleichung

folgt unmittelbar aus der Integralform der Schwarzschen Ungleichung[1])

$$\left(\int_a^b x(t)\, y(t)\, dt\right)^2 \leqq \int_a^b x^2(t)\, dt \cdot \int_a^b y^2(t)\, dt.$$

9. Wir betrachten die Menge aller beschränkten Zahlenfolgen von reellen Zahlen $x = (x_1, x_2, \ldots, x_n, \ldots)$. Wenn wir

$$\varrho(x, y) = \sup_k |y_k - x_k| \tag{11}$$

setzen, erhalten wir einen metrischen Raum, den wir mit m bezeichnen. Die Gültigkeit der Axiome 1 bis 3 ist trivial.

10. Die Menge aller n-Tupel von reellen Zahlen mit dem Abstand

$$\varrho_p(x, y) = \left(\sum_{k=1}^{n} |y_k - x_k|^p\right)^{1/p}, \tag{12}$$

wobei $p \geqq 1$ eine beliebige feste Zahl ist, bildet einen metrischen Raum, den wir mit $\boldsymbol{R}_p{}^n$ bezeichnen. Die Gültigkeit der Axiome 1 und 2 ist hier wiederum offensichtlich. Wir beweisen, daß auch Axiom 3 erfüllt ist. Dazu seien

$$x = (x_1, \ldots, x_n), \qquad y = (y_1, \ldots, y_n), \qquad z = (z_1, \ldots, z_n)$$

drei Punkte aus $\boldsymbol{R}_p{}^n$. Wir setzen

$$y_k - x_k = a_k, \qquad z_k - y_k = b_k;$$

dann nimmt die Ungleichung

$$\varrho_p(x, z) \leqq \varrho_p(x, y) + \varrho_p(y, z),$$

deren Richtigkeit wir nachweisen müssen, die Form

$$\left(\sum_{k=1}^{n} |a_k + b_k|^p\right)^{1/p} \leqq \left(\sum_{k=1}^{n} |a_k|^p\right)^{1/p} + \left(\sum_{k=1}^{n} |b_k|^p\right)^{1/p} \tag{13}$$

an. Das ist die sogenannte *Minkowskische Ungleichung*. Für $p = 1$ gilt die Minkowskische Ungleichung trivialerweise (der Betrag einer Summe ist kleiner oder gleich der Summe der Beträge); daher werden wir annehmen, daß $p > 1$ ist.[2])

---

[1]) Diese Ungleichung erhält man z. B. aus der leicht zu beweisenden Identität

$$\left(\int_a^b x(t)\, y(t)\, dt\right)^2 = \int_a^b x^2(t)\, dt \int_a^b y^2(t)\, dt - \frac{1}{2} \int_a^b \int_a^b [x(s)\, y(t) - y(s)\, x(t)]^2\, ds\, dt.$$

[2]) Für $p < 1$ gilt die Minkowskische Ungleichung nicht. Wollten wir also den Raum $\boldsymbol{R}_p{}^n$ für $p < 1$ betrachten, dann wäre in diesem Raum die Dreiecksungleichung nicht erfüllt.

Der Beweis der Ungleichung (13) für $p > 1$ beruht auf der sogenannten *Hölderschen Ungleichung*

$$\sum_{k=1}^{n} |a_k b_k| \leqq \left(\sum_{k=1}^{n} |a_k|^p\right)^{1/p} \left(\sum_{k=1}^{n} |b_k|^q\right)^{1/q}, \tag{14}$$

wobei die Zahlen $p > 1$ und $q > 1$ die Bedingung

$$\frac{1}{p} + \frac{1}{q} = 1, \quad \text{d. h.} \quad q = \frac{p}{p-1} \tag{15}$$

erfüllen.

Wir stellen fest, daß die Ungleichung (14) homogen ist. Das heißt, wenn sie für zwei beliebige Vektoren

$$a = (a_1, \ldots, a_n) \quad \text{und} \quad b = (b_1, \ldots, b_n)$$

gültig ist, dann gilt sie auch für die Vektoren λa und μb, wobei λ und μ beliebige Zahlen sind. Daher genügt es, Ungleichung (14) für den Fall

$$\sum_{k=1}^{n} |a_k|^p = \sum_{k=1}^{n} |b_k|^q = 1 \tag{16}$$

zu beweisen.

Es sei also die Bedingung (16) erfüllt; wir zeigen dann

$$\sum_{k=1}^{n} |a_k b_k| \leqq 1. \tag{17}$$

Dazu betrachten wir in der ξ, η-Ebene die Kurve, die durch die Gleichung $\eta = \xi^{p-1}$ ($\xi > 0$) oder, was genau dasselbe ist, durch die Gleichung $\xi = \eta^{q-1}$ (Abb. 7) bestimmt ist. Aus Abb. 7 geht hervor, daß bei beliebiger Wahl der reellen Zahlen a und b

$$S_1 + S_2 \geqq a \cdot b$$

ist. Die Berechnung der Flächen S_1 und S_2 ergibt

$$S_1 = \int_0^a \xi^{p-1}\, d\xi = \frac{a^p}{p}, \qquad S_2 = \int_0^b \eta^{q-1}\, d\eta = \frac{b^q}{q}.$$

Abb. 7

Also gilt die Ungleichung

$$ab \leqq \frac{a^p}{p} + \frac{b^q}{q}$$

für beliebige reelle Zahlen a und b. Wenn wir hier a durch $|a_k|$ und b durch $|b_k|$ ersetzen und über k von 1 bis n summieren, erhalten wir unter Berücksichtigung von (15) und (16)

$$\sum_{k=1}^{n} |a_k b_k| \leqq 1.$$

Damit ist Ungleichung (17) und folglich auch die allgemeinere Ungleichung (14) gezeigt. Für $p = 2$ geht die Höldersche Ungleichung (14) in die Schwarzsche Ungleichung (4) über.

Wir kommen jetzt zum Beweis der Minkowskischen Ungleichung. Dazu betrachten wir die Identität

$$(|a| + |b|)^p = (|a| + |b|)^{p-1} |a| + (|a| + |b|)^{p-1} |b|.$$

Wenn wir in dieser Gleichung a durch a_k und b durch b_k ersetzen und über k von 1 bis n summieren, erhalten wir

$$\sum_{k=1}^{n} (|a_k| + |b_k|)^p = \sum_{k=1}^{n} (|a_k| + |b_k|)^{p-1} |a_k| + \sum_{k=1}^{n} (|a_k| + |b_k|)^{p-1} |b_k|.$$

Wenden wir jetzt auf jede der beiden rechts stehenden Summen die Höldersche Ungleichung an und berücksichtigen, daß $(p - 1)\, q = p$ ist, so erhalten wir

$$\sum_{k=1}^{n} (|a_k| + |b_k|)^p \leqq \left(\sum_{k=1}^{n} (|a_k| + |b_k|)^p\right)^{1/q} \left(\left[\sum_{k=1}^{n} |a_k|^p\right]^{1/p} + \left[\sum_{k=1}^{n} |b_k|^p\right]^{1/p}\right).$$

Dividieren wir noch beide Seiten dieser Gleichung durch

$$\left(\sum_{k=1}^{n} (|a_k| + |b_k|)^p\right)^{1/q},$$

so ergibt sich

$$\left(\sum_{k=1}^{n} (|a_k| + |b_k|)^p\right)^{1/p} \leqq \left(\sum_{k=1}^{n} |a_k|^p\right)^{1/p} + \left(\sum_{k=1}^{n} |b_k|^p\right)^{1/p},$$

woraus sofort Ungleichung (13) folgt. Damit ist die Dreiecksungleichung in $\boldsymbol{R}_n^p$ bewiesen.

Die in diesem Beispiel betrachtete Metrik ϱ_p ist für $p = 2$ die euklidische Metrik (Beispiel 3) und für $p = 1$ die Metrik von Beispiel 4. Man kann zeigen, daß die

Metrik

$$\varrho_0(x, y) = \max_{1 \leqq k \leqq n} |y_k - x_k|,$$

die in Beispiel 5 eingeführt wurde, der Grenzfall der Metriken $\varrho_p(x, y)$ ist; es gilt nämlich

$$\varrho_0(x, y) = \lim_{p \to \infty} \left(\sum_{k=1}^{\infty} |y_k - x_k|^p \right)^{1/p}.$$

Aus der oben gezeigten Ungleichung

$$ab \leqq \frac{a^p}{p} + \frac{b^q}{q} \quad \left(\frac{1}{p} + \frac{1}{q} = 1 \right)$$

leitet man auch leicht die *Höldersche Ungleichung für Integrale*,

$$\int_a^b |x(t)\, y(t)|\, dt \leqq \left(\int_a^b |x(t)|^p\, dt \right)^{1/p} \left(\int_a^b |y(t)|^q\, dt \right)^{1/q},$$

ab, die für beliebige Funktionen $x(t)$ und $y(t)$ richtig ist, für welche die rechts stehenden Integrale einen Sinn haben. Daraus wiederum ergibt sich die *Minkowskische Ungleichung für Integrale:*

$$\left(\int_a^b |x(t) + y(t)|^p\, dt \right)^{1/p} \leqq \left(\int_a^b |x(t)|^p\, dt \right)^{1/p} + \left(\int_a^b |y(t)|^p\, dt \right)^{1/p}.$$

11. Wir erwähnen noch ein interessantes Beispiel eines metrischen Raumes. Seine Elemente sind alle Folgen

$$x = (x_1, x_2, \ldots, x_n, \ldots)$$

reeller Zahlen mit der Eigenschaft

$$\sum_{k=1}^{\infty} |x_k|^p < \infty,$$

wobei $p \geqq 1$ eine feste Zahl ist, und der Abstand wird durch die Formel

$$\varrho(x, y) = \left(\sum_{k=1}^{\infty} |y_k - x_k|^p \right)^{1/p} \tag{18}$$

definiert. Diesen metrischen Raum bezeichnen wir mit l_p.

Nach der Minkowskischen Ungleichung (13) gilt für beliebiges n

$$\left(\sum_{k=1}^{n} |y_k - x_k|^p \right)^{1/p} \leqq \left(\sum_{k=1}^{n} |x_k|^p \right)^{1/p} + \left(\sum_{k=1}^{n} |y_k|^p \right)^{1/p}.$$

Weil nach Voraussetzung die Reihen

$$\sum_{k=1}^{\infty} |x_k|^p \quad \text{und} \quad \sum_{k=1}^{\infty} |y_k|^p$$

konvergieren, erhalten wir durch den Grenzübergang $n \to \infty$

$$\left(\sum_{k=1}^{\infty} |y_k - x_k|^p\right)^{1/p} \leqq \left(\sum_{k=1}^{\infty} |x_k|^p\right)^{1/p} + \left(\sum_{k=1}^{\infty} |y_k|^p\right)^{1/p} < \infty. \tag{19}$$

Auf diese Weise ist gezeigt, daß Formel (18), die den Abstand in l_p definiert, für beliebige $x, y \in l_p$ einen Sinn besitzt. Gleichzeitig zeigt Ungleichung (19), daß in l_p die Dreiecksungleichung gilt. Die restlichen Axiome sind offensichtlich erfüllt.

Eine beliebige Anzahl weiterer Beispiele ergibt sich nach folgendem Verfahren. Es sei $R = (X, \varrho)$ ein metrischer Raum und M eine beliebige Teilmenge von X. Dann bildet M mit derselben Funktion $\varrho(x, y)$, die wir jetzt nur für x und y aus M betrachten, ebenfalls einen metrischen Raum; er heißt *Teilraum* des Raumes R.

2.1.2. Stetige Abbildungen metrischer Räume. Isometrie. Es seien X und Y zwei metrische Räume und f eine Abbildung des Raumes X in Y. Es wird also jedem Element $x \in X$ ein Element $y = f(x)$ aus Y zugeordnet. Diese Abbildung heißt *stetig* im Punkt $x_0 \in X$, wenn für beliebiges $\varepsilon > 0$ ein $\delta > 0$ existiert, so daß für alle $x \in X$ mit

$$\varrho(x, x_0) < \delta$$

die Ungleichung

$$\varrho_1\big(f(x), f(x_0)\big) < \varepsilon$$

erfüllt ist (hier ist ϱ die Metrik in X und ϱ_1 die Metrik in Y). Wenn die Abbildung f in allen Punkten des Raumes X stetig ist, sagt man, daß f auf X *stetig* ist. Sind X und Y Mengen reeller Zahlen, d. h., ist f eine reellwertige Funktion, die auf einer Teilmenge X der Zahlengeraden definiert ist, dann stimmt die angegebene Definition der Stetigkeit einer Abbildung mit der aus der elementaren Analysis wohlbekannten Definition der Stetigkeit von Funktionen überein.

Analog kann man eine stetige Funktion (Abbildung) f von mehreren Veränderlichen $x_1 \in X_1, \ldots, x_n \in X_n$ (wobei $X_1, \ldots, X_n$ metrische Räume sind) mit Werten in einem metrischen Raum Y definieren.

Wir bemerken in diesem Zusammenhang, daß auch der Abstand $\varrho(x, y)$ stetig ist, wenn man ihn als Funktion der Veränderlichen x und y aus X betrachtet. Das folgt sofort aus der Ungleichung

$$|\varrho(x, y) - \varrho(x_0, y_0)| \leqq \varrho(x_0, x) + \varrho(y_0, y),$$

die leicht aus der Dreiecksungleichung abgeleitet werden kann.

Ist die Abbildung $f\colon X \to Y$ umkehrbar eindeutig von X auf Y, so existiert die Umkehrabbildung $x = f^{-1}(y)$ des Raumes Y auf den Raum X. Sind in diesem Fall sowohl die Abbildung f als auch die Umkehrabbildung f^{-1} stetig, dann heißt f *homöomorphe Abbildung* oder *Homöomorphismus*. Dabei heißen die Räume X und Y, zwischen denen man den Homöomorphismus herstellen kann, zueinander *homöomorph*. Als Beispiel homöomorpher Räume können die ganze reelle Achse $(-\infty, \infty)$ und ein offenes Intervall, zum Beispiel $(-1, 1)$, dienen. In diesem Fall wird durch die Formel

$$y = \frac{2}{\pi} \arctan x$$

ein Homöomorphismus hergestellt.

Einen wichtigen Spezialfall homöomorpher Abbildungen bilden die sogenannten isometrischen Abbildungen. Man sagt, daß eine Bijektion f zwischen zwei metrischen Räumen $R = (X, \varrho)$ und $R' = (Y, \varrho')$ eine *Isometrie* ist, wenn

$$\varrho(x_1, x_2) = \varrho'\big(f(x_1), f(x_2)\big)$$

für beliebige $x_1, x_2 \in R$ ist. Räume R und R', zwischen denen man eine isometrische Abbildung herstellen kann, heißen *isometrisch*.

Isometrie der Räume R und R' bedeutet, daß die metrischen Beziehungen zwischen ihren Elementen ein und dieselben sind. Verschieden kann nur die Natur ihrer Elemente sein, was jedoch vom Standpunkt der Theorie der metrischen Räume unwesentlich ist. Im weiteren werden wir zueinander isometrische Räume einfach als identisch ansehen.

Auf die hier eingeführten Begriffe (Stetigkeit, Homöomorphismus) werden wir von einem allgemeineren Standpunkt aus am Ende von 2.5. zurückkommen.

2.2. Konvergenz. Offene und abgeschlossene Mengen

2.2.1. Grenzwert. Abschluß. Wir führen jetzt einige Begriffe aus der Theorie der metrischen Räume ein. Diese Begriffe werden wir im folgenden wiederholt benutzen.

Als *offene Kugel* $B(x_0, r)$ im metrischen Raum R bezeichnen wir die Gesamtheit aller Punkte $x \in R$, die der Bedingung

$$\varrho(x, x_0) < r$$

genügen. Der Punkt x_0 heißt *Mittelpunkt* und die Zahl r *Radius* dieser Kugel.

Als *abgeschlossene Kugel* $B[x_0, r]$ bezeichnen wir die Menge aller Punkte $x \in R$, die die Bedingung

$$\varrho(x, x_0) \leqq r$$

erfüllen.

Eine offene Kugel mit dem Mittelpunkt x_0 und dem Radius ε werden wir auch *ε-Umgebung* des Punktes x_0 nennen und mit dem Symbol $O_\varepsilon(x_0)$ bezeichnen.

Aufgabe. Man gebe ein Beispiel eines metrischen Raumes mit zwei Kugeln $B(x, \varrho_1)$, $B(y, \varrho_2)$ an, für die $\varrho_1 > \varrho_2$ ist und trotzdem $B(x, \varrho_1) \subset B(y, \varrho_2)$ gilt.

Ein Punkt $x \in R$ heißt *Berührungspunkt* einer Menge $M \subset R$, wenn jede Umgebung von x wenigstens einen Punkt aus M enthält. Die Menge aller Berührungspunkte der Menge M wird mit $[M]$ bezeichnet und heißt *Abschluß* von M. Auf diese Weise haben wir für Mengen eines metrischen Raumes die *Operation der Abschließung* erklärt, den Übergang von der Menge M zu ihrem Abschluß $[M]$.

Satz 1. *Die Abschließung besitzt folgende Eigenschaften:*

1. $M \subset [M$,

2. $[[M]] = [M]$,

3. $[M_1] \subset [M_2]$ *falls* $M_1 \subset M_2$,

4. $[M_1 \cup M_2] = [M_1] \cup [M_2]$.

Beweis. Die erste Behauptung gilt trivialerweise, weil jeder Punkt, der zu M gehört, auch ein Berührungspunkt von M ist.

Wir beweisen die zweite Eigenschaft. Dazu sei $x \in [[M]]$. Dann gibt es in einer beliebigen Umgebung $O_\varepsilon(x)$ von x einen Punkt $x_1 \in [M]$. Wir setzen $\varepsilon_1 = \varepsilon - \varrho(x, x_1)$ und betrachten die Kugel $O_{\varepsilon_1}(x_1)$. Diese Kugel liegt ganz im Innern der Kugel $O_\varepsilon(x)$. Denn für $z \in O_{\varepsilon_1}(x_1)$ gilt $\varrho(z, x_1) < \varepsilon_1$, und wegen $\varrho(x, x_1) = \varepsilon - \varepsilon_1$ folgt nach der Dreiecksungleichung

$$\varrho(z, x) < \varepsilon_1 + (\varepsilon - \varepsilon_1) = \varepsilon,$$

d. h. $z \in O_\varepsilon(x)$. Weil nun $x_1 \in [M]$ ist, gibt es in $O_{\varepsilon_1}(x_1)$ einen Punkt $x_2 \in M$. Dann ist aber auch $x_2 \in O_\varepsilon(x)$. Da $O_\varepsilon(x)$ eine beliebige Umgebung des Punktes x war, folgt $x \in [M]$. Damit ist die zweite Behauptung gezeigt.

Abschließend beweisen wir noch die vierte Eigenschaft, die dritte ist offensichtlich erfüllt.

Wenn $x \in [M_1 \cup M_2]$ ist, liegt x wenigstens in einer der Mengen $[M_1]$ oder $[M_2]$, d. h., es ist

$$[M_1 \cup M_2] \subset [M_1] \cup [M_2].$$

Wegen $M_1 \subset M_1 \cup M_2$ und $M_2 \subset M_1 \cup M_2$ folgt auch die umgekehrte Inklusion aus Eigenschaft 3. Damit ist der Satz vollständig bewiesen.

Ein Punkt $x \in R$ heißt *Häufungspunkt* der Menge $M \subset R$, wenn jede Umgebung von x unendlich viele Punkte aus M enthält.

Ein Häufungspunkt kann zu M gehören, braucht aber nicht zu M zu gehören. Ist z. B. M die Menge aller rationalen Zahlen aus dem abgeschlossenen Intervall $[0, 1]$, so ist jeder Punkt dieses Intervalls Häufungspunkt von M.

Ein Punkt x, der zu M gehört, heißt *isolierter Punkt* dieser Menge, wenn es in einer hinreichend kleinen Umgebung $O_\varepsilon(x)$ von x keine von x verschiedenen Punkte aus M gibt.

Der Leser möge folgende Behauptung als Übung beweisen:

Jeder Berührungspunkt einer Menge M ist entweder Häufungspunkt oder isolierter Punkt dieser Menge.

Hieraus kann man folgern, daß der Abschluß $[M]$ im allgemeinen aus drei Arten von Punkten besteht:

1. aus isolierten Punkten der Menge M;
2. aus Häufungspunkten der Menge M, die zu M gehören;
3. aus Häufungspunkten der Menge M, die nicht zu M gehören.

Somit erhält man den Abschluß $[M]$ durch Vereinigung der Menge M mit allen ihren Häufungspunkten.

2.2.2. Konvergenz. Es sei $x_1, x_2, \ldots$ eine Folge von Punkten im metrischen Raum R. Man sagt, daß diese Folge *gegen den Punkt x konvergiert*, wenn jede Umgebung $O_\varepsilon(x)$ des Punktes x alle Punkte x_n von einem gewissen Index an enthält, d. h., wenn es zu jedem $\varepsilon > 0$ eine Zahl N_ε gibt, so daß $O_\varepsilon(x)$ alle Punkte x_n mit $n > N_\varepsilon$ enthält. Der Punkt x heißt *Grenzwert* der Folge $\{x_n\}$.

Diese Definition kann man offensichtlich auch noch in folgender Weise formulieren. Die Folge $\{x_n\}$ konvergiert gegen x, wenn

$$\lim_{n\to\infty} \varrho(x, x_n) = 0.$$

Aus der Definition des Grenzwertes folgt unmittelbar, daß erstens eine Folge nicht zwei verschiedene Grenzwerte besitzen kann und daß zweitens auch jede Teilfolge einer Folge $\{x_n\}$, die gegen den Punkt x konvergiert, gegen denselben Punkt konvergiert.

Der folgende Satz stellt den Zusammenhang zwischen den Begriffen des Berührungspunktes und des Häufungspunktes her.

Satz 2. *Ein Punkt x ist genau dann Berührungspunkt einer Menge M, wenn eine Folge $\{x_n\}$ von Punkten aus M existiert, die gegen x konvergiert.*

Beweis. Wir zeigen die Notwendigkeit der Bedingung, die Hinlänglichkeit ist trivial. Wenn x Berührungspunkt der Menge M ist, liegt in jeder Umgebung $O_{1/n}(x)$ von x wenigstens ein Punkt $x_n \in M$. Diese Punkte bilden eine Folge, die gegen x konvergiert.

Ist x Häufungspunkt der Menge M, so kann man die Punkte $x_n \in O_{1/n} \cap M$, die verschiedenen n entsprechen, paarweise verschieden wählen. *Somit ist der Punkt x dann und nur dann Häufungspunkt von M, wenn in M eine Folge von paarweise verschiedenen Punkten existiert, die gegen x konvergiert.*

Den Begriff der Stetigkeit einer Abbildung eines metrischen Raumes X in einen metrischen Raum Y, der in 2.1. eingeführt wurde, kann man jetzt mit Hilfe der Konvergenz von Folgen ausdrücken. Und zwar ist eine Abbildung $y = f(x)$ stetig im Punkt x_0, wenn für jede Folge $\{x_n\}$, die gegen x_0 konvergiert, die Folge $\{y_n = f(x_n)\}$ gegen $y_0 = f(x_0)$ konvergiert. Der Beweis der Gleichwertigkeit dieser Definition mit der in 2.1. angegebenen unterscheidet sich durch nichts vom Beweis der Gleichwertigkeit der beiden Stetigkeitsdefinitionen (in der „ε-δ-Sprache“ und in der „Folgensprache“) für Zahlenfunktionen und kann dem Leser überlassen werden.

2.2.3. Dichte Teilmengen. Es seien A und B zwei Mengen in einem metrischen Raum R. Die Menge A heißt *dicht* in B, wenn $[A] \supset B$ ist. Insbesondere heißt die Menge A *überall dicht* (im Raum R), wenn ihr Abschluß $[A]$ mit dem ganzen Raum R übereinstimmt. Zum Beispiel ist die Menge der rationalen Zahlen überall dicht auf der Zahlengeraden. Die Menge A heißt *nirgends dicht,* wenn sie in keiner Kugel dicht ist, d. h., wenn in jeder Kugel $B \subset R$ eine andere Kugel B' enthalten ist, die mit M keinen gemeinsamen Punkt besitzt.

Beispiele von Räumen, die eine abzählbare überall dichte Menge besitzen. Räume, in denen es eine abzählbare überall dichte Menge gibt, heißen *separabel.* Wir wollen die in 2.1. angeführten Beispiele unter diesem Gesichtspunkt betrachten.

1. Der „diskrete“ Raum, der in 2.1., Beispiel 1, beschrieben wurde, enthält dann und nur dann eine abzählbare überall dichte Menge, wenn er selbst nur aus abzählbar vielen Punkten besteht. Das liegt daran, daß der Abschluß $[M]$ einer beliebigen Menge M in diesem Raum mit M übereinstimmt.

Alle Räume, die in 2.1., Beispiel 2 bis 8, aufgezählt wurden, enthalten abzählbare überall dichte Mengen. Wir geben in jedem dieser Räume je eine solche Menge an und empfehlen dem Leser dringend, die detaillierten Beweise durchzuführen.

2. Auf der reellen Achse $\boldsymbol{R}^1$ die rationalen Zahlen.

3.—5. Im n-dimensionalen euklidischen Raum $\boldsymbol{R}^n$ und in den Räumen $\boldsymbol{R}_1{}^n$, $\boldsymbol{R}_0{}^n$ die Gesamtheit aller Vektoren mit rationalen Koordinaten.

6. Im Raum $C[a, b]$ die Gesamtheit aller Polynome mit rationalen Koeffizienten.

7. Im Raum l_2 die Gesamtheit aller Folgen mit rationalen Folgengliedern, von denen nur endlich viele von Null verschieden sind.

8. Im Raum $C^2[a, b]$ die Gesamtheit aller Polynome mit rationalen Koeffizienten.

Hingegen ist der *Raum m aller beschränkten Folgen* (2.1., Beispiel 9) nicht separabel. Zum Beweis betrachten wir alle Folgen, die aus Nullen und Einsen bestehen. Sie bilden eine Menge von der Mächtigkeit des Kontinuums (weil man zwischen ihnen und den Teilmengen der natürlichen Zahlen eine eineindeutige Beziehung herstellen kann). Der Abstand zwischen zwei solchen Punkten, die durch Formel (11) aus 2.1. definiert wird, ist gleich 1. Um jeden dieser Punkte legen wir eine offene Kugel vom Radius 1/2. Diese Kugeln schneiden sich nicht. Wenn nun eine Menge in m überall dicht ist, dann muß jede der konstruierten Kugeln wenigstens einen Punkt aus dieser Menge enthalten, und diese kann folglich nicht abzählbar sein.

2.2.4. Offene und abgeschlossene Mengen. Wir betrachten zwei wichtige Typen von Mengen im metrischen Raum, und zwar die offenen und die abgeschlossenen Mengen.

Eine Menge M in einem metrischen Raum R heißt *abgeschlossen*, wenn sie mit ihrem Abschluß übereinstimmt: $[M] = M$. Eine Menge ist also abgeschlossen, wenn sie alle ihre Häufungspunkte enthält.

Nach Satz 1 ist der Abschluß einer Menge M abgeschlossen. Aus demselben Satz folgt, daß $[M]$ die kleinste abgeschlossene Menge ist, die M enthält. (Man beweise das!)

Beispiele

1. Jedes abgeschlossene Intervall $[a, b]$ auf der Zahlengeraden ist eine abgeschlossene Menge.

2. Eine abgeschlossene Kugel stellt eine abgeschlossene Menge dar. Insbesondere ist im Raum $C[a, b]$ die Menge der Funktionen abgeschlossen, die die Bedingung $|f(t)| \leq K$ erfüllen.

3. In $C[a, b]$ ist die Menge aller Funktionen mit $|f(t)| < K$ (offene Kugel) nicht abgeschlossen, ihr Abschluß ist die abgeschlossene Kugel, die aus allen Funktionen mit $|f(t)| \leq K$ besteht.

4. In jedem metrischen Raum R sind die leere Menge $\emptyset$ und der ganze Raum R abgeschlossen.

5. Jede Menge, die aus endlich vielen Elementen besteht, ist abgeschlossen.

Im folgenden Satz formulieren wir grundlegende Eigenschaften abgeschlossener Mengen.

Satz 3. *Der Durchschnitt beliebig vieler und die Vereinigung endlich vieler abgeschlossener Mengen sind abgeschlossene Mengen.*

Beweis. Es sei $F = \bigcap F_\alpha$ der Durchschnitt der abgeschlossenen Mengen F_α, und es sei x ein Häufungspunkt von F. Das bedeutet, daß eine beliebige Umgebung $O_\varepsilon(x)$ von x unendlich viele Punkte aus F enthält. Dann enthält aber $O_\varepsilon(x)$ erst recht unendlich viele Punkte aus jedem F_α, und folglich gehört der Punkt x zu jedem F_α, weil jedes F_α abgeschlossen ist. Somit folgt $x \in F = \bigcap F_\alpha$, d. h., F ist abgeschlossen.

Es sei jetzt F die Vereinigung endlich vieler abgeschlossener Mengen: $F = \bigcup_{i=1}^{n} F_i$, und es sei x ein Punkt, der nicht zu F gehört. Wir werden zeigen, daß x dann nicht Häufungspunkt von F sein kann. Der Punkt x gehört nämlich zu keiner der abgeschlossenen Mengen F_i und ist folglich auch nicht Häufungspunkt einer dieser Mengen. Daher kann man für jedes i eine Umgebung $O_{\varepsilon_i}(x)$ des Punktes x finden, die höchstens endlich viele Punkte aus F_i enthält. Nehmen wir nun die kleinste der Umgebungen $O_{\varepsilon_1}(x), \ldots, O_{\varepsilon_n}(x)$, so erhalten wir eine Umgebung $O_\varepsilon(x)$ des Punktes x, die höchstens endlich viele Punkte aus F enthält.

Wenn also der Punkt x nicht zu F gehört, dann kann er nicht Häufungspunkt von F sein, d. h., F ist abgeschlossen. Damit ist der Satz bewiesen.

Ein Punkt x heißt *innerer Punkt* der Menge M, wenn eine Umgebung $O_\varepsilon(x)$ dieses Punktes existiert, die ganz in M enthalten ist.

Eine Menge, die nur aus inneren Punkten besteht, heißt *offen*.

Beispiele

6. Das offene Intervall (a, b) auf der Zahlengeraden $\boldsymbol{R}^1$ ist eine offene Menge; wenn nämlich $a < \alpha < b$ ist, dann liegt $O_\varepsilon(\alpha)$ mit $\varepsilon = \min(\alpha - a, b - \alpha)$ ganz im Intervall (a, b).

7. Eine offene Kugel $B(a, r)$ in einem beliebigen metrischen Raum R ist eine offene Menge. Denn für jedes $x \in B(a, r)$ ist $\varepsilon = r - \varrho(a, x) > 0$ und $B(x, \varepsilon) \subset B(a, r)$.

8. Die Menge aller auf $[a, b]$ stetigen Funktionen, die der Bedingung $f(t) < g(t)$ genügen, wobei $g(t)$ eine feste stetige Funktion ist, bildet eine offene Teilmenge des Raumes $C[a, b]$.

Satz 4. *Eine Menge M ist genau dann offen, wenn ihr Komplement $R \setminus M$ abgeschlossen ist.*

Beweis. Wenn M offen ist, besitzt jeder Punkt x aus M eine Umgebung, die ganz zu M gehört, d. h. die keinen gemeinsamen Punkt mit $R \setminus M$ hat. Ein Punkt, der nicht zu $R \setminus M$ gehört, kann demnach kein Berührungspunkt von $R \setminus M$ sein. Also ist $R \setminus M$ abgeschlossen. Ist umgekehrt $R \setminus M$ abgeschlossen, so besitzt jeder Punkt aus M eine Umgebung, die ganz in M liegt, d. h., M ist offen.

Weil die leere Menge und der ganze Raum R abgeschlossen sind und gleichzeitig die Komplemente voneinander bilden, *sind die leere Menge und der ganze Raum R ebenfalls offen.*

Aus Satz 3 und aus dem Dualitätsprinzip (der Durchschnitt der Komplemente ist gleich dem Komplement der Vereinigung, die Vereinigung der Komplemente ist gleich dem Komplement des Durchschnitts; vgl. S. 19) ergibt sich folgender wichtiger Satz, der zu Satz 3 dual ist.

Satz 3'. *Die Vereinigung beliebig vieler und der Durchschnitt endlich vieler offener Mengen sind offene Mengen.*

2.2.5. Offene und abgeschlossene Mengen auf der Zahlengeraden. Die Struktur offener und abgeschlossener Mengen kann in manchen metrischen Räumen sehr kompliziert sein. Das trifft auch auf offene und abgeschlossene Mengen im zwei- oder dreidimensionalen euklidischen Raum zu. Nur im eindimensionalen Fall, d. h. auf der Zahlengeraden, bereitet die vollständige Beschreibung aller offenen Mengen (und folglich auch aller abgeschlossenen) keine Mühe. Sie ergibt sich aus folgendem Satz.

Satz 5. *Jede offene Menge auf der Zahlengeraden läßt sich als Vereinigung endlich oder abzählbar vieler paarweise disjunkter offener Intervalle*[1]) *darstellen.*

Beweis. G sei eine offene Menge auf der reellen Achse. Für die Punkte aus G führen wir eine Äquivalenzrelation ein, indem wir $x \sim y$ setzen, wenn ein offenes Intervall I mit $x, y \in I \subset G$ existiert. Offensichtlich ist diese Relation reflexiv und symmetrisch. Sie ist auch transitiv. Ist nämlich $x \sim y$ und $y \sim z$, dann existieren offene Intervalle I_1 und I_2 mit

$$x, y \in I_1 \subset G \quad \text{und} \quad y, z \in I_2 \subset G.$$

Die Vereinigung $I = I_1 \cup I_2$ ist wegen $y \in I_1 \cap I_2$ offenbar ebenfalls ein Intervall, für das

$$x, z \in I \subset G$$

gilt. Das bedeutet aber $x \sim z$. Folglich zerfällt G in disjunkte Klassen I_τ von äquivalenten Punkten,

$$G = \bigcup I_\tau .$$

Wir werden zeigen, daß jedes I_τ ein offenes Intervall (a, b) ist, wobei $a = \inf I_\tau$, $b = \sup I_\tau$ ist. Die Inklusion $I_\tau \subset (a, b)$ ist trivial. Gilt andererseits $x, y \in I_\tau$, so ist nach Definition von I_τ das Intervall (x, y) in I_τ enthalten. Nun gibt es in beliebiger Nähe rechts von a und in beliebiger Nähe links von b Punkte aus I_τ. Daher enthält I_τ jedes Intervall (a', b'), dessen Endpunkte zu (a, b) gehören, woraus $I_\tau = (a, b)$ folgt. Das System dieser disjunkten Intervalle I_τ ist höchstens abzählbar, denn wenn wir aus jedem dieser Intervalle in beliebiger Weise einen rationalen Punkt auswählen, erhalten wir eine umkehrbar eindeutige Beziehung zwischen diesen Intervallen und einer gewissen Teilmenge der rationalen Zahlen. Damit ist der Satz bewiesen.

Jede abgeschlossene Menge ist das Komplement einer offenen Menge. Somit erhält man jede abgeschlossene Zahlenmenge, indem man der reellen Achse endlich oder abzählbar viele offene Intervalle entnimmt.

Die einfachsten Beispiele abgeschlossener Mengen sind abgeschlossene Intervalle, einzelne Punkte und endliche Vereinigungen solcher Mengen. Wir wollen auch noch

[1]) Die Mengen der Form $(-\infty, \infty)$, (α, ∞) $(-\infty, \beta)$ sehen wir dabei ebenfalls als Intervalle an.

ein komplizierteres Beispiel einer abgeschlossenen Menge auf der Zahlengeraden betrachten, und zwar die sogenannte *Cantorsche Menge.*

Es sei F_0 das abgeschlossene Intervall $[0, 1]$. Wir entnehmen daraus das offene Intervall $(1/3, 2/3)$ und bezeichnen die verbleibende abgeschlossene Menge mit F_1. Daraufhin nehmen wir aus F_1 die offenen Intervalle $(1/9, 2/9)$ und $(7/9, 8/9)$ heraus und bezeichnen die verbleibende abgeschlossene Menge (die aus vier abgeschlossenen Intervallen besteht) mit F_2. Aus jedem dieser vier Intervalle nehmen wir wieder

Abb. 8

das mittlere offene Intervall der Länge $(1/3)^3$ heraus usw. (Abb. 8). Wenn wir diesen Prozeß fortsetzen, erhalten wir eine abnehmende Folge abgeschlossener Mengen F_n. Setzen wir

$$F = \bigcap_{n=0}^{\infty} F_n,$$

dann ist F als Durchschnitt von abgeschlossenen Mengen eine abgeschlossene Menge. Diese Menge erhält man aus dem abgeschlossenen Intervall $[0, 1]$ durch die Entnahme von abzählbar vielen offenen Intervallen.

Wir wollen die Struktur der Menge F untersuchen. Zu F gehören offensichtlich die Punkte

$$0, 1, \frac{1}{3}, \frac{2}{3}, \frac{1}{9}, \frac{2}{9}, \frac{7}{9}, \frac{8}{9}, \ldots, \tag{1}$$

die Endpunkte der herausgenommenen offenen Intervalle. Jedoch sind das nicht alle Punkte von F. Die Punkte des Intervalls $[0, 1]$, die zur Menge F gehören, kann man auf folgende Weise charakterisieren. Wir schreiben jede der Zahlen x, $0 \leqq x \leqq 1$, im triadischen System:

$$x = \frac{a_1}{3} + \frac{a_2}{3^2} + \cdots + \frac{a_n}{3^n} + \cdots,$$

wobei die Zahlen a_n die Werte 0, 1, 2 annehmen können. Wie auch im Fall von dekadischen Brüchen gestatten einige Zahlen zwei Schreibweisen, z. B.

$$\frac{1}{3} = \frac{1}{3} + \frac{0}{3^2} + \cdots + \frac{0}{3^n} + \cdots = \frac{0}{3} + \frac{2}{3^2} + \cdots + \frac{2}{3^n} + \cdots$$

Man kann leicht nachprüfen, daß die Zahl x, $0 \leqq x \leqq 1$, genau dann zur Menge F gehört, wenn man sie wenigstens auf eine Weise so als triadischen Bruch schreiben kann, daß in der Folge $a_1, a_2, \ldots, a_n, \ldots$ niemals die Eins vorkommt. Somit kann man jedem Punkt $x \in F$ eine Folge

$$a_1, a_2, \ldots, a_n, \ldots \tag{2}$$

zuordnen, wobei $a_n = 0$ oder $a_n = 2$ ist. Die Gesamtheit dieser Folgen bildet eine Menge von der Mächtigkeit des Kontinuums. Davon kann man sich überzeugen, indem man jeder Folge (2) die Folge

$$b_1, b_2, \ldots, b_n, \ldots \tag{2'}$$

mit $b_n = 0$ für $a_n = 0$ und $b_n = 1$ für $a_n = 2$ zuordnet. Die Folge (2') kann man nämlich als Darstellung einer reellen Zahl y, $0 \leqq y \leqq 1$, als dyadischen Bruch betrachten. Somit erhalten wir eine Abbildung der Menge F auf das ganze Intervall $[0, 1]$. Hieraus folgt, daß F die Mächtigkeit des Kontinuums besitzt.[1]) Weil die Menge der Punkte (1) abzählbar ist, können diese Punkte nicht ganz F ausschöpfen.

Aufgaben

1. Es ist direkt zu beweisen, daß der Punkt $^1/_4$ zur Menge F gehört, aber kein Endpunkt eines der herausgenommenen Intervalle ist.

Hinweis. Der Punkt $^1/_4$ teilt das Intervall $[0, 1]$ im Verhältnis 1:3. Das Intervall $[0, ^1/_3]$, das nach der ersten Herausnahme verbleibt, teilt er ebenfalls im Verhältnis 1:3 usw.

Die Punkte (1) nennen wir Punkte *erster Art* der Menge F, die verbleibenden Punkte nennen wir Punkte *zweiter Art.*

2. Es ist zu beweisen, daß die Punkte erster Art in F eine überall dichte Menge bilden.

3. Es ist zu zeigen, daß die Zahlen der Form $t_1 + t_2$, $t_1, t_2 \in F$, das ganze Intervall $[0, 2]$ ausfüllen.

Wir haben gezeigt, daß die Menge F die Mächtigkeit des Kontinuums besitzt, d. h. ebensoviele Punkte wie das Intervall $[0, 1]$.

Es ist interessant, das folgende Resultat mit dieser Tatsache zu vergleichen: Die Summe $\frac{1}{3} + \frac{2}{9} + \frac{4}{27} + \cdots$ der Längen aller herausgenommenen Intervalle ist genau gleich 1.

Ergänzende Bemerkungen

(1) Es sei M eine Menge im metrischen Raum R und x ein Punkt dieses Raumes. Dann wird die Zahl

$$\varrho(x, M) = \inf_{a \in M} \varrho(x, a)$$

Abstand des Punktes x von der Menge M genannt.

[1]) Die aufgestellte Beziehung zwischen F und dem Intervall $[0, 1]$ ist eindeutig, aber nicht eineindeutig (deswegen, weil manche Zahlen des Intervalls $[0, 1]$ durch verschiedene dyadische Brüche dargestellt werden können). Hieraus folgt, daß die Mächtigkeit von F nicht kleiner als die Mächtigkeit des Kontinuums ist. Andererseits ist F eine Teilmenge des Intervalls $[0, 1]$, und daher kann die Mächtigkeit von F nicht größer als die Mächtigkeit des Kontinuums sein.

Ist $x \in M$, so ist $\varrho(x, M) = 0$, jedoch aus $\varrho(x, M) = 0$ folgt nicht, daß $x \in M$ ist. Aus der Definition des Berührungspunktes erhalten wir unmittelbar, daß $\varrho(x, M) = 0$ dann und nur dann gilt, wenn x Berührungspunkt der Menge M ist.

Somit kann man die Abschließung einer Menge auch dadurch charakterisieren, daß alle die Punkte zur Menge hinzugenommen werden, deren Abstand zur Menge gleich Null ist.

(2) Analog wird die Entfernung zwischen zwei Mengen definiert. Wenn A und B zwei Mengen im metrischen Raum R sind, so ist

$$\varrho(A, B) = \inf_{\substack{a \in A \\ b \in B}} \varrho(a, b).$$

Wenn $A \cap B \neq \emptyset$, so ist $\varrho(A, B) = 0$; das Umgekehrte ist allgemein nicht richtig.

(3) K sei eine feste positive Zahl, und M_K sei die Menge aller Funktionen f aus $C[a, b]$, die der Lipschitzbedingung

$$|f(t_1) - f(t_2)| \leqq K\,|t_2 - t_1|, \qquad t_1, t_2 \in [a, b],$$

genügen. Die Menge M_K ist abgeschlossen. Sie stimmt mit dem Abschluß der Menge aller der auf $[a, b]$ differenzierbaren Funktionen überein, für die $|f'(t)| \leqq K$ ist.

(4) Die Menge $M = \bigcup_K M_K$ derjenigen Funktionen, die die Lipschitzbedingung für jedes K erfüllen, ist nicht abgeschlossen. Ihr Abschluß ist ganz $C[a, b]$.

(5) Eine offene Menge G im n-dimensionalen euklidischen Raum heißt *zusammenhängend*, wenn je zwei Punkte $x, y \in G$ durch einen Polygonzug verbunden werden können, der ganz in G liegt. Zum Beispiel ist das Innere des Kreises $x^2 + y^2 < 1$ eine zusammenhängende Menge. Dagegen ist die Vereinigung der beiden Kreise

$$x^2 + y^2 < 1 \quad \text{und} \quad (x - 2)^2 + y^2 < 1$$

keine zusammenhängende Menge (obwohl diese Kreise einen gemeinsamen Berührungspunkt haben!). Eine offene Teilmenge H der offenen Menge G heißt *Komponente* von G, wenn sie zusammenhängend ist und in keiner größeren zusammenhängenden offenen Teilmenge von G enthalten ist. Wir führen in G eine Äquivalenzrelation ein. Es sei $x \sim y$, wenn eine offene zusammenhängende Teilmenge H von G existiert, die x und y enthält,

$$x, y \in H \subset G.$$

Wie im Fall der reellen Achse prüft man leicht die Transitivität nach, und daher zerfällt G in disjunkte Klassen $G = \bigcup I$. Diese Klassen sind die offenen Komponenten von G. Ihre Zahl ist höchstens abzählbar.

Im Fall $n = 1$, d. h. auf der Geraden, ist jede zusammenhängende offene Menge ein offenes Intervall (zu den offenen Intervallen werden auch die unendlichen Intervalle $(-\infty, a)$, (b, ∞) und $(-\infty, \infty)$ gerechnet). Somit besteht Satz 5, der eine Aussage über die Struktur offener Mengen auf der Zahlengeraden macht, aus zwei Behauptungen:

1. Jede offene Menge auf der Zahlengeraden ist die Vereinigung endlich vieler oder abzählbar vieler Komponenten.

2. Eine zusammenhängende offene Menge auf der Zahlengeraden ist ein offenes Intervall.

Die erste Behauptung ist auch für Mengen im n-dimensionalen euklidischen Raum richtig (und gestattet weiteste Verallgemeinerungen), aber die zweite Behauptung bezieht sich speziell auf die Zahlengerade.

2.3. Vollständige metrische Räume

2.3.1. Definition und Beispiele vollständiger metrischer Räume. Schon in den Anfangsgründen der Analysis erkennt man, was für eine wichtige Rolle die Vollständigkeit der Zahlengeraden in der Analysis spielt, d. h. die Tatsache, daß jede Fundamentalfolge reeller Zahlen gegen einen Grenzwert konvergiert. Die Zahlengerade dient als einfachstes Beispiel der sogenannten vollständigen metrischen Räume, deren grundlegende Eigenschaften wir in 2.3. untersuchen werden.

Eine Folge $\{x_n\}$ von Punkten eines metrischen Raumes R nennen wir *Fundamentalfolge*, wenn sie dem Cauchyschen Kriterium genügt, d. h., wenn für jedes $\varepsilon > 0$ eine Zahl N_ε existiert, so daß $\varrho(x_{n'}, x_{n''}) < \varepsilon$ ist für alle $n' > N_\varepsilon$, $n'' > N_\varepsilon$.

Definition 1. Wenn im Raum R jede Fundamentalfolge konvergiert, dann heißt dieser Raum *vollständig*.

Beispiele. Alle Räume, die in 2.1. betrachtet wurden, sind mit Ausnahme des in Beispiel 8 angegebenen Raumes vollständig. Das werden wir jetzt zeigen.

1. Fundamentalfolgen im Raum isolierter Punkte (2.1., Beispiel 1) sind nur die stationären Folgen, d. h. Folgen, in denen von einem gewissen Index an alle Folgenglieder gleich sind. Jede solche Folge konvergiert natürlich, d. h., dieser Raum ist vollständig.

2. Die Vollständigkeit des euklidischen Raumes $\boldsymbol{R}^1$, der Gesamtheit aller reellen Zahlen, ist aus der elementaren Analysis bekannt.

3. Die Vollständigkeit des euklidischen Raumes $\boldsymbol{R}^n$ folgt unmittelbar aus der Vollständigkeit des $\boldsymbol{R}^1$. Um das einzusehen, betrachten wir eine Fundamentalfolge $\{x^{(p)}\}$ von Punkten des $\boldsymbol{R}^n$. Dann gibt es zu jedem $\varepsilon > 0$ ein $N = N_\varepsilon$, so daß

$$\sum_{k=1}^{n} (x_k^{(p)} - x_k^{(q)})^2 < \varepsilon^2$$

ist für alle p, q, die größer als N sind. Hierbei ist $x^{(p)} = \{x_1^{(p)}, \ldots, x_n^{(p)}\}$. Daraus erhalten wir für jedes $k = 1, 2, \ldots, n$ eine entsprechende Ungleichung für die Koordinaten $x_k^{(p)}$:

$$|x_k^{(p)} - x_k^{(q)}| < \varepsilon$$

für alle $p, q > N$. Also ist $\{x_k^{(p)}\}$ eine Fundamentalfolge im Raum $\boldsymbol{R}^1$. Setzen wir nun

$$x_k = \lim_{p\to\infty} x_k^{(p)} \quad \text{und} \quad x = (x_1, x_2, \ldots, x_n),$$

dann ist offenbar

$$\lim_{p\to\infty} x^{(p)} = x.$$

4.—5. Die Vollständigkeit der Räume $\boldsymbol{R}_0{}^n$ und $\boldsymbol{R}_1{}^n$ wird völlig analog bewiesen.

6. Wir zeigen jetzt die Vollständigkeit des Raumes $C[a, b]$. Dazu sei $\{x_n(t)\}$ eine Fundamentalfolge in $C[a, b]$. Das bedeutet, daß zu jedem $\varepsilon > 0$ eine Zahl N existiert, so daß

$$|x_n(t) - x_m(t)| < \varepsilon$$

ist für $n, m > N$ und für alle t, $a \leqq t \leqq b$. Hieraus folgt, daß die Folge $\{x_n(t)\}$ gleichmäßig konvergiert. Bekanntlich ist ihr Grenzwert $x(t)$ dann eine stetige Funktion. Lassen wir nun m in der vorhergehenden Ungleichung gegen Unendlich streben, so erhalten wir

$$|x_n(t) - x(t)| \leqq \varepsilon$$

für alle t und alle $n > N$. Das heißt aber, daß $\{x_n(t)\}$ im Sinne der Metrik des Raumes $C[a, b]$ gegen $x(t)$ konvergiert.

7. Um die Vollständigkeit des Raumes l_2 zu zeigen, betrachten wir eine Fundamentalfolge $\{x^{(n)}\}$ in l_2. Dann gibt es zu jedem $\varepsilon > 0$ ein N, so daß

$$\varrho^2(x^{(n)}, x^{(m)}) = \sum_{k=1}^{\infty} (x_k{}^{(n)} - x_k{}^{(m)})^2 < \varepsilon \tag{1}$$

für $n, m > N$ gilt. Hierbei ist

$$x^{(n)} = (x_1{}^{(n)}, x_2{}^{(n)}, \ldots, x_k{}^{(n)}, \ldots).$$

Aus (1) folgt für jedes k die Ungleichung

$$(x_k{}^{(n)} - x_k{}^{(m)})^2 < \varepsilon,$$

d. h., für jedes k ist die reelle Zahlenfolge $\{x_k{}^{(n)}\}$ Fundamentalfolge und daher konvergent. Wir setzen nun $x_k = \lim\limits_{n\to\infty} x_k{}^{(n)}$ und bezeichnen die Folge $(x_1, x_2, \ldots, x_k, \ldots)$ mit x. Dann ist zu zeigen:

a) $\sum\limits_{k=1}^{\infty} x_k{}^2 < \infty$, d. h. $x \in l_2$,

b) $\lim\limits_{n\to\infty} \varrho(x^{(n)}, x) = 0$.

Aus Ungleichung (1) folgt, daß für beliebiges festes M die Ungleichung

$$\sum_{k=1}^{M} (x_k{}^{(n)} - x_k{}^{(m)})^2 < \varepsilon$$

gilt. In dieser Summe gibt es nur endlich viele Summanden, und wir können bei festgehaltenem n den Grenzübergang $m \to \infty$ ausführen. Wir erhalten dann

$$\sum_{k=1}^{M} (x_k{}^{(n)} - x_k)^2 \leqq \varepsilon.$$

Diese Ungleichung ist für jedes M richtig. Daher können wir durch den Grenzübergang $M \to \infty$ wieder zur unendlichen Reihe übergehen und erhalten

$$\sum_{k=1}^{\infty} (x_k^{(n)} - x_k)^2 \leqq \varepsilon. \tag{2}$$

Aus der Konvergenz der Reihen $\sum_{k=1}^{\infty} (x_k^{(n)})^2$ und $\sum_{k=1}^{\infty} (x_k^{(n)} - x_k)^2$ folgt nun (auf Grund der elementaren Ungleichung $(a+b)^2 \leqq 2(a^2+b^2)$) die Konvergenz der Reihe $\sum_{k=1}^{\infty} x_k^2$, d. h., Behauptung a) ist gezeigt. Da ε beliebig klein war, ergibt sich aus Ungleichung (2) ferner

$$\lim_{n\to\infty} \varrho(x^{(n)}, x) = \lim_{n\to\infty} \sqrt{\sum_{k=1}^{\infty} (x_k^{(n)} - x_k)^2} = 0,$$

d. h. $x^{(n)} \to x$ in der Metrik des l_2. Damit ist auch Behauptung b) bewiesen.

8. Man überzeugt sich leicht davon, daß der Raum $C^2[a, b]$ nicht vollständig ist. Dazu betrachten wir z. B. die Folge der stetigen Funktionen

$$\varphi_n(t) = \begin{cases} -1 & \text{für} \quad -1 \leqq t \leqq -\dfrac{1}{n}, \\ nt & \text{für} \quad -\dfrac{1}{n} \leqq t \leqq \dfrac{1}{n}, \\ 1 & \text{für} \quad \dfrac{1}{n} \leqq t \leqq 1. \end{cases}$$

Sie ist Fundamentalfolge in $C^2[-1, 1]$ wegen

$$\int_{-1}^{1} \bigl(\varphi_n(t) - \varphi_m(t)\bigr)^2\, dt \leqq \frac{2}{\min(n, m)}.$$

Jedoch konvergiert sie gegen keine Funktion aus $C^2[-1, 1]$. Um das zu zeigen, betrachten wir eine Funktion f aus $C^2[-1, 1]$ und die Sprungfunktion ψ, die für $t < 0$ gleich -1 und für $t \geqq 0$ gleich $+1$ ist.

Nach der Minkowskischen Ungleichung für Integrale (die offenbar auch für stückweise stetige Funktionen gilt) erhalten wir

$$\left(\int_{-1}^{1} \bigl(f(t) - \psi(t)\bigr)^2\, dt\right)^{1/2}$$
$$\leqq \left(\int_{-1}^{1} \bigl(f(t) - \varphi_n(t)\bigr)^2\, dt\right)^{1/2} + \left(\int_{-1}^{1} \bigl(\varphi_n(t) - \psi(t)\bigr)^2\, dt\right)^{1/2}.$$

Wegen der Stetigkeit der Funktion f ist das Integral auf der linken Seite von Null verschieden. Ferner ist klar, daß

$$\lim_{n\to\infty} \int_{-1}^{1} \left(\varphi_n(t) - \psi(t)\right)^2 dt = 0$$

ist. Daher kann $\int_{-1}^{1} \left(f(t) - \varphi_n(t)\right)^2 dt$ für $n \to \infty$ nicht gegen Null streben.

Aufgabe. Es ist zu beweisen, daß der Raum aller beschränkten Folgen (2.1., Beispiel 9) vollständig ist.

2.3.2. Der Cantorsche Durchschnittssatz. In der elementaren Analysis wird oft das sogenannte *Intervallschachtelungsprinzip* ausgenutzt. In der Theorie der metrischen Räume spielt der folgende Satz, der sogenannte *Cantorsche Durchschnittssatz,* eine ähnliche Rolle.

Satz 1. *Ein metrischer Raum ist genau dann vollständig, wenn in ihm jede Folge von ineinandergeschachtelten abgeschlossenen Kugeln, deren Radien gegen Null streben, einen nichtleeren Durchschnitt besitzt.*

Beweis. Zunächst zeigen wir die Notwendigkeit. Der Raum R sei vollständig, und es sei $B_1, B_2, B_3, \ldots$ eine Folge von ineinandergeschachtelten abgeschlossenen Kugeln. Der Radius der Kugel B_n sei r_n und der Mittelpunkt x_n. Die Folge der Mittelpunkte $\{x_n\}$ ist dann eine Fundamentalfolge, denn es gilt $\varrho(x_n, x_m) < r_n$ für $m > n$ und $r_n \to 0$ für $n \to \infty$. Also existiert $\lim\limits_{n\to\infty} x_n = x$, da R vollständig ist. Wenn wir nun

$$x = \lim_{n\to\infty} x_n$$

setzen, dann ist $x \in \bigcap\limits_n B_n$. Die Kugel B_n enthält nämlich alle Punkte der Folge $\{x_k\}$, eventuell mit Ausnahme der Punkte $x_1, x_2, \ldots, x_{n-1}$. Somit ist x Häufungspunkt jeder Kugel B_n. Da aber B_n abgeschlossen ist, folgt $x \in B_n$ für alle n.

Die Bedingung ist auch hinreichend. Es sei $\{x_n\}$ eine Fundamentalfolge, und wir wollen zeigen, daß sie einen Grenzwert besitzt. Weil $\{x_n\}$ Fundamentalfolge ist, können wir einen Punkt x_{n_1} auswählen, so daß $\varrho(x_n, x_{n_1}) < 1/2$ ist für alle $n \geqq n_1$. Wir nehmen den Punkt x_{n_1} als Zentrum einer abgeschlossenen Kugel vom Radius 1 und bezeichnen diese Kugel mit B_1. Danach wählen wir x_{n_2} aus $\{x_n\}$ so, daß $n_2 > n_1$ und $\varrho(x_n, x_{n_2}) < 1/2^2$ ist für $n \geqq n_2$. B_2 sei die abgeschlossene Kugel mit dem Mittelpunkt x_{n_2} und dem Radius 1/2. Sind allgemein die Punkte $x_{n_1}, x_{n_2}, \ldots, x_{n_k}$ bereits gewählt $(n_1 < n_2 < \cdots < n_k)$, so wählen wir einen Punkt $x_{n_{k+1}}$ so, daß $n_{k+1} > n_k$ ist und $\varrho(x_n, x_{n_{k+1}}) < 1/2^{k+1}$ für alle $n \geqq n_{k+1}$. Als B_{k+1} nehmen wir die abgeschlossene Kugel um $x_{n_{k+1}}$ mit dem Radius $1/2^k$. Setzen wir diese Konstruktion fort, so erhalten

wir eine Folge von ineinandergeschachtelten abgeschlossenen Kugeln B_k, wobei die Kugel B_k den Radius $1/2^{k-1}$ besitzt. Diese Folge von Kugeln hat nach Voraussetzung einen gemeinsamen Punkt, wir bezeichnen ihn mit x. Es ist klar, daß dieser Punkt x der Grenzwert der Teilfolge $\{x_{n_k}\}$ ist. Aber wenn eine Fundamentalfolge eine gegen den Punkt x konvergierende Teilfolge enthält, dann konvergiert sie selbst gegen eben diesen Grenzwert. Somit gilt $x = \lim_{n\to\infty} x_n$. Damit ist der Satz bewiesen.

Aufgaben

1. Es ist zu zeigen, daß der Durchschnitt der ineinandergeschachtelten abgeschlossenen Kugeln im vorhergehenden Satz aus einem Punkt besteht.

2. *Durchmesser* der Menge M im metrischen Raum heißt die Zahl

$$d(M) = \sup_{x,y\in M} \varrho(x, y).$$

Es ist zu beweisen, daß in einem vollständigen metrischen Raum jede Folge von ineinandergeschachtelten nichtleeren abgeschlossenen Mengen, deren Durchmesser gegen Null streben, einen nichtleeren Durchschnitt besitzt.

3. Es ist ein Beispiel eines vollständigen metrischen Raumes und einer Folge von darinliegenden ineinandergeschachtelten abgeschlossenen Kugeln anzugeben, die einen leeren Durchschnitt besitzen.

4. Es ist zu beweisen, daß ein Teilraum eines vollständigen metrischen Raumes R dann und nur dann vollständig ist, wenn er in R abgeschlossen ist.

2.3.3. Der Satz von Baire. In der Theorie der vollständigen metrischen Räume spielt der folgende Satz eine fundamentale Rolle.

Satz 2 (Baire). *Ein vollständiger metrischer Raum R kann nicht als Vereinigung von abzählbar vielen nirgends dichten Mengen dargestellt werden.*

Beweis. Wir nehmen das Gegenteil an. Es sei also $R = \bigcup_{n=1}^{\infty} M_n$, wobei jede der Mengen M_n nirgends dicht ist. S_0 sei irgendeine abgeschlossene Kugel vom Radius 1. Die nirgends dichte Menge M_1 ist nicht dicht in S_0, also gibt es eine abgeschlossene Kugel S_1 mit einem Radius kleiner als 1/2, so daß $S_1 \subset S_0$ und $S_1 \cap M_1 = \emptyset$ ist. Da die Menge M_2 in S_1 nicht dicht ist, gibt es nach derselben Überlegung in S_1 eine abgeschlossene Kugel S_2, deren Radius kleiner als 1/3 ist und für die $S_2 \cap M_2 = \emptyset$ gilt, usw. Wir erhalten auf diese Weise eine Folge $\{S_n\}$ von ineinandergeschachtelten Kugeln, deren Radien gegen Null streben, wobei $S_n \cap M_n = \emptyset$ ist. Nach Satz 1 enthält der Durchschnitt $\bigcap_{n=1}^{\infty} S_n$ einen Punkt x. Dieser Punkt gehört nach Konstruktion zu keiner der Mengen M_n. Somit gilt $x \notin \bigcup_{n=1}^{\infty} M_n$, d. h., es ist $R \neq \bigcup_{n=1}^{\infty} M_n$, im Widerspruch zur Voraussetzung.

Insbesondere ist also *jeder vollständige metrische Raum ohne isolierte Punkte überabzählbar*. Denn in einem solchen Raum ist jeder Punkt nirgends dicht.

2.3.4. Die Vervollständigung von Räumen. Wenn ein Raum R nicht vollständig ist, dann kann man ihn immer auf irgendeine (und im wesentlichen einzige) Weise in einen vollständigen Raum einbetten.

Definition 2. Es sei R ein metrischer Raum. Ein vollständiger metrischer Raum R^* heißt dann *Vervollständigung* des Raumes R, wenn:

1. R Teilraum des Raumes R^* ist;

2. R überall dicht ist in R^*, d. h. $[R] = R^*$. ($[R]$ ist der Abschluß von R in R^*.)

Zum Beispiel ist der Raum aller reellen Zahlen die Vervollständigung des Raumes der rationalen Zahlen.

Satz 3. *Jeder metrische Raum R besitzt eine Vervollständigung, und diese ist bis auf eine Isometrie, die die Punkte von R invariant läßt, eindeutig bestimmt.*

Beweis. Wir beginnen mit der Eindeutigkeit. Wir müssen zeigen, daß zu zwei Vervollständigungen R^* und R^{**} des Raumes R eine umkehrbar eindeutige Abbildung φ des Raumes R^* auf den Raum R^{**} existiert, so daß

1. $\varphi(x) = x$ für alle $x \in R$ gilt;

2. $\varrho_1(x^*, y^*) = \varrho_2(x^{**}, y^{**})$ für $x^* \leftrightarrow x^{**}$ und $y^* \leftrightarrow y^{**}$ ist; dabei ist ϱ_1 der Abstand in R^* und ϱ_2 der Abstand in R^{**}.

Die Abbildung φ wird folgendermaßen definiert. Es sei x^* ein beliebiger Punkt aus R^*. Dann existiert nach Definition der Vervollständigung eine Folge $\{x_n\}$ von Punkten aus R, die gegen x^* konvergiert. Die Punkte x_n liegen aber auch in R^{**}. Da R^{**} vollständig ist, konvergiert $\{x_n\}$ in R^{**} gegen einen Punkt x^{**}. Es ist klar, daß x^{**} nicht von der Auswahl der Folge $\{x_n\}$ abhängt, die gegen den Punkt x^* konvergiert. Setzen wir nun $\varphi(x^*) = x^{**}$, so ist φ die gesuchte isometrische Abbildung. Denn nach Konstruktion gilt erstens $\varphi(x) = x$ für alle $x \in R$. Weiter folgt aus

$$\{x_n\} \to x^* \text{ in } R^* \quad \text{und} \quad \{x_n\} \to x^{**} \text{ in } R^{**},$$

$$\{y_n\} \to y^* \text{ in } R^* \quad \text{und} \quad \{y_n\} \to y^{**} \text{ in } R^{**}$$

wegen der Stetigkeit des Abstandes die Beziehung

$$\varrho_1(x^*, y^*) = \lim_{n\to\infty} \varrho_1(x_n, y_n) = \lim_{n\to\infty} \varrho(x_n, y_n)$$

und analog

$$\varrho_2(x^{**}, y^{**}) = \lim_{n\to\infty} \varrho_2(x_n, y_n) = \lim_{n\to\infty} \varrho(x_n, y_n).$$

Folglich gilt auch

$$\varrho_1(x^*, y^*) = \varrho_2(x^{**}, y^{**}).$$

Wir zeigen jetzt die Existenz der Vervollständigung. Die Beweisidee ist dieselbe wie in der Cantorschen Theorie der reellen Zahlen. Die Situation ist hier sogar ein-

facher als bei den reellen Zahlen, weil dort für die neu eingeführten Objekte, die irrationalen Zahlen, noch alle arithmetischen Operationen definiert werden mußten.

Es sei also R ein beliebiger metrischer Raum. Wir nennen zwei Fundamentalfolgen $\{x_n\}$ und $\{x_n'\}$ aus R äquivalent (Bezeichnung $\{x_n\} \sim \{x_n'\}$), wenn $\lim_{n\to\infty} (x_n, x_n') = 0$ ist. Diese Bezeichnung wird dadurch gerechtfertigt, daß die so erklärte Relation reflexiv, symmetrisch und transitiv ist. Daher zerfällt die Menge aller Fundamentalfolgen, die man aus Punkten des Raumes R bilden kann, in Klassen zueinander äquivalenter Fundamentalfolgen. Wir definieren jetzt den Raum R^*. Als dessen Punkte nehmen wir alle Äquivalenzklassen von Fundamentalfolgen, und den Abstand zwischen ihnen erklären wir folgendermaßen. Es seien x^* und y^* zwei solche Klassen. Aus jeder dieser Klassen wählen wir je einen Repräsentanten, d. h. je eine Fundamentalfolge $\{x_n\}$ und $\{y_n\}$, und setzen[1])

$$\varrho(x^*, y^*) = \lim_{n\to\infty} \varrho(x_n, y_n). \tag{3}$$

Wir weisen zunächst die Korrektheit dieser Abstandsdefinition nach, d. h., wir zeigen, daß der Grenzwert (3) existiert und nicht von der Auswahl der Repräsentanten $\{x_n\} \in x^*$ und $\{y_n\} \in y^*$ abhängt. Aus der Ungleichung

$$|\varrho(x_n, y_n) - \varrho(x_m, y_m)| \leqq \varrho(x_n, x_m) + \varrho(y_n, y_m) \tag{4}$$

erhalten wir für hinreichend große n und m

$$|\varrho(x_n, y_n) - \varrho(x_m, y_m)| < \varepsilon,$$

da $\{x_n\}$ und $\{y_n\}$ Fundamentalfolgen sind. Somit erfüllt die Folge $s_n = \varrho(x_n, y_n)$ von reellen Zahlen das Cauchysche Kriterium und besitzt folglich einen Grenzwert. Dieser Grenzwert hängt nicht von der Auswahl $\{x_n\} \in x^*$ und $\{y_n\} \in y^*$ ab. Denn wenn

$$\{x_n\}, \{x_n'\} \in x^* \quad \text{und} \quad \{y_n\}, \{y_n'\} \in y^*$$

ist, ergibt sich vollkommen analog zu (4)

$$|\varrho(x_n, y_n) - \varrho(x_n', y_n')| \leqq \varrho(x_n, x_n') + \varrho(y_n, y_n'),$$

und wegen $\{x_n\} \sim \{x_n'\}$ und $\{y_n\} \sim \{y_n'\}$ folgt hieraus

$$\lim_{n\to\infty} \varrho(x_n, y_n) = \lim_{n\to\infty} \varrho(x_n', y_n').$$

Wir weisen nun nach, daß R^* die Axiome eines metrischen Raumes erfüllt.

Die Gültigkeit von Axiom 1 ergibt sich unmittelbar aus der Definition der Äquivalenz von Fundamentalfolgen. Axiom 2 ist offensichtlich erfüllt. Es muß also noch die Gültigkeit der Dreiecksungleichung überprüft werden. Da die Dreiecksungleichung

[1]) Um die Bezeichnungsweise nicht zu erschweren, bezeichnen wir den Abstand in R^* mit demselben Symbol ϱ wie den Abstand im Ausgangsraum R.

im Ausgangsraum R gilt, ist

$$\varrho(x_n, z_n) \leqq \varrho(x_n, y_n) + \varrho(y_n, z_n).$$

Durch den Grenzübergang $n \to \infty$ erhalten wir daraus

$$\lim_{n\to\infty} \varrho(x_n, z_n) \leqq \lim_{n\to\infty} \varrho(x_n, y_n) + \lim_{n\to\infty} \varrho(y_n, z_n),$$

d. h.

$$\varrho(x^*, z^*) \leqq \varrho(x^*, y^*) + \varrho(y^*, z^*).$$

Jetzt zeigen wir, daß man R als Teilraum des Raumes R^* betrachten kann. Jedem Punkt $x \in R$ entspricht eine gewisse Klasse äquivalenter Fundamentalfolgen, und zwar die Gesamtheit aller Folgen, die gegen den Punkt x konvergieren. Diese Klasse ist nicht leer, da sie die stationäre Folge enthält, deren Folgenglieder alle gleich x sind. Außerdem gilt für $x, y \in R$

$$\varrho(x, y) = \lim_{n\to\infty} \varrho(x_n, y_n),$$

wenn $\lim_{n\to\infty} x_n = x$ und $\lim_{n\to\infty} y_n = y$ ist. Folglich bilden wir R isometrisch in den Raum R^* ab, wenn wir jedem $x \in R$ die Klasse aller gegen x konvergierenden Fundamentalfolgen zuordnen.

Im folgenden werden wir nicht mehr zwischen dem Raum R selbst und seinem Bild in R^* unterscheiden und betrachten R als Teilraum von R^*.

Nun zeigen wir, daß R überall dicht ist in R^*. Dazu sei x^* irgendein Punkt aus R^* und $\varepsilon > 0$ beliebig. Wir wählen in x^* einen Repräsentanten, d. h. eine gewisse Fundamentalfolge $\{x_n\}$. N sei jetzt so groß, daß $\varrho(x_n, x_m) < \varepsilon$ ist für alle $n, m > N$. Dann haben wir

$$\varrho(x_n, x^*) = \lim_{m\to\infty} \varrho(x_n, x_m) \leqq \varepsilon$$

für $n > N$, d. h., eine beliebige Umgebung des Punktes x^* enthält einen Punkt aus R. Somit ist der Abschluß von R in R^* ganz R^*.

Es bleibt die Vollständigkeit von R^* zu zeigen. Wir bemerken zunächst, daß nach Konstruktion von R^* eine beliebige Fundamentalfolge

$$x_1, x_2, \ldots, x_n, \ldots$$

von Punkten aus R in R^* gegen einen gewissen Punkt konvergiert, und zwar gegen den Punkt $x^* \in R$, der durch eben diese Folge definiert wird. Weil R dicht in R^* ist, folgt nun aber, daß man jeder Fundamentalfolge $x_1^*, x_2^*, \ldots x_n^*, \ldots$ von Punkten aus R^* eine zu ihr äquivalente Folge $x_1, x_2, \ldots, x_n, \ldots$ von Punkten aus R konstruieren kann. Dazu genügt es, für x_n einen beliebigen Punkt aus R mit $\varrho(x_n, x_n^*) < 1/n$ zu nehmen. Die so konstruierte Folge ist Fundamentalfolge in R und konvergiert nach Definition von R^* gegen einen Punkt $x^* \in R^*$. Dann konvergiert aber auch die Folge $\{x_n^*\}$ gegen x^*. Damit ist der Satz vollständig bewiesen.

2.4. Das Prinzip der kontrahierenden Abbildung und seine Anwendung

2.4.1. Das Prinzip der kontrahierenden Abbildung. Eine Reihe von Fragen, die mit der Existenz und Eindeutigkeit von Lösungen einiger Gleichungstypen (z. B. Differentialgleichungen) zusammenhängen, formuliert man mitunter auch als Frage nach der Existenz und der Eindeutigkeit eines Fixpunktes einer gewissen Abbildung, die einen metrischen Raum in sich überführt. Unter verschiedenen Kriterien der Existenz und Eindeutigkeit eines Fixpunktes ist das sogenannte *Prinzip der kontrahierenden Abbildung* für Abbildungen dieser Art eines der einfachsten und zugleich wichtigsten.

Es sei R ein metrischer Raum. Eine Abbildung A des Raumes R in sich heißt *kontrahierende Abbildung* oder kürzer *Kontraktion*, wenn eine Zahl $\alpha < 1$ existiert, so daß für zwei beliebige Punkte $x, y \in R$ die Ungleichung

$$\varrho(Ax, Ay) \leqq \alpha\varrho(x, y) \tag{1}$$

erfüllt ist. Jede kontrahierende Abbildung ist offenbar stetig. Denn aus $x_n \to x$ folgt nach (1) auch $Ax_n \to Ax$.

Ein Punkt x heißt *Fixpunkt* der Abbildung A, wenn $Ax = x$ gilt. Fixpunkte sind also die Lösungen der Gleichung $Ax = x$.

Satz 1 (*Prinzip der kontrahierenden Abbildung*). *Jede kontrahierende Abbildung in einem vollständigen metrischen Raum besitzt genau einen Fixpunkt.*

Beweis. Es sei x_0 ein beliebiger Punkt in R. Dann setzen wir $x_1 = Ax_0, x_2 = Ax_1 = A^2x_0$ usw. und allgemein $x_n = Ax_{n-1} = A^nx_0$. Wir zeigen nun, daß die Folge $\{x_n\}$ eine Fundamentalfolge ist. Wenn wir etwa annehmen, daß $m \geqq n$ ist, gilt

$$\begin{aligned}\varrho(x_n, x_m) &= \varrho(A^nx_0, A^mx_0) \leqq \alpha^n\varrho(x_0, x_{m-n}) \\ &\leqq \alpha^n\{\varrho(x_0, x_1) + \varrho(x_1, x_2) + \cdots + \varrho(x_{m-n-1}, x_{m-n})\} \\ &\leqq \alpha^n\varrho(x_0, x_1)\{1 + \alpha + \alpha^2 + \cdots + \alpha^{m-n-1}\} \\ &\leqq \alpha^n\varrho(x_0, x_1)\frac{1}{1-\alpha}.\end{aligned}$$

Wegen $\alpha < 1$ wird diese Größe bei hinreichend großem n beliebig klein. Somit ist die Folge $\{x_n\}$ Fundamentalfolge, und wegen der Vollständigkeit von R besitzt sie einen Grenzwert x,

$$x = \lim_{n\to\infty} x_n.$$

Auf Grund der Stetigkeit der Abbildung A folgt dann

$$Ax = A\left(\lim_{n\to\infty} x_n\right) = \lim_{n\to\infty} Ax_n = \lim_{n\to\infty} x_{n+1} = x.$$

Damit ist die Existenz eines Fixpunktes bewiesen.

Wir weisen jetzt die Eindeutigkeit nach. Wenn $Ax = x$, $Ay = y$ ist, nimmt (1) die Form $\varrho(x, y) \leqq \alpha\varrho(x, y)$ an, und wegen $\alpha < 1$ folgt hieraus $\varrho(x, y) = 0$, d. h. $x = y$.

Aufgabe. Es ist an einem Beispiel zu zeigen, daß eine Abbildung A, die für alle $x \neq y$ der Bedingung $\varrho(Ax, Ay) < \varrho(x, y)$ genügt, keinen einzigen Fixpunkt zu besitzen braucht.

2.4.2. Einfache Anwendungen des Prinzips der kontrahierenden Abbildung. Das Prinzip der kontrahierenden Abbildung kann man zum Beweis von Existenz- und Unitätssätzen für Lösungen verschiedener Gleichungstypen anwenden. Abgesehen vom Existenz- und Eindeutigkeitsbeweis für die Lösung der Gleichung $Ax = x$, liefert das Prinzip der kontrahierenden Abbildung auch eine praktikable Methode zur näherungsweisen Berechnung dieser Lösung (*Methode der sukzessiven Approximation*). Dazu betrachten wir die folgenden einfachen Beispiele.

1. Die Funktion f sei auf dem abgeschlossenen Intervall $[a, b]$ definiert, erfülle die Lipschitzbedingung

$$|f(x_2) - f(x_1)| \leqq K\,|x_2 - x_1|$$

mit einer Konstanten $K < 1$ und bilde das Intervall $[a, b]$ in sich ab. Dann ist f eine kontrahierende Abbildung, und nach dem bewiesenen Satz konvergiert die Folge $x_0, x_1 = f(x_0), x_2 = f(x_1), \ldots$ gegen die einzige Lösung der Gleichung $x = f(x)$.

Insbesondere ist die Kontraktionsbedingung erfüllt, wenn die Funktion auf dem Intervall $[a, b]$ eine Ableitung $f'(x)$ besitzt mit

$$|f'(x)| \leqq K < 1.$$

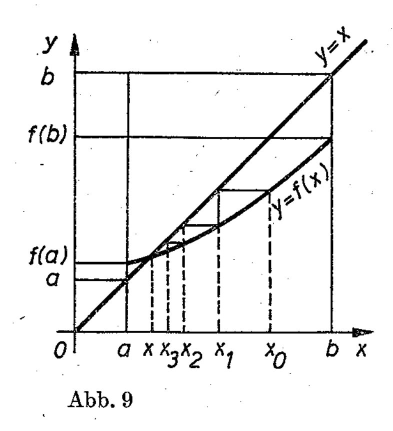


Abb. 9

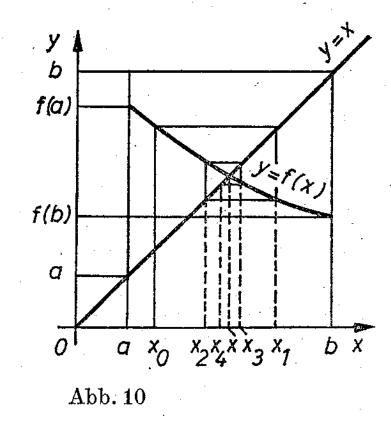


Abb. 10

In Abb. 9 und 10 ist der Verlauf der sukzessiven Approximation im Fall $0 < f'(x) < 1$ und im Fall $-1 < f'(x) < 0$ dargestellt.

Wir wollen uns jetzt mit Gleichungen der Form $F(x) = 0$ beschäftigen, wobei $F(a) < 0$, $F(b) > 0$ und $0 < K_1 \leqq F'(x) \leqq K_2$ auf $[a, b]$ ist. Wir führen dazu die Funktion $f(x) = x - \lambda F(x)$ ein und suchen eine Lösung der Gleichung $x = f(x)$, die mit $F(x) = 0$ gleichwertig ist. Wegen $f'(x) = 1 - \lambda F'(x)$ ist

$$1 - \lambda K_2 \leqq f'(x) \leqq 1 - \lambda K_1,$$

und die Zahl λ kann nun leicht so gewählt werden, daß man die Methode der sukzessiven Approximation anwenden darf. Das ist eine weitverbreitete Lösungsmethode.

2. Wir betrachten die Abbildung A des n-dimensionalen Raumes in sich, die durch das lineare Gleichungssystem

$$y_i = \sum_{j=1}^{n} a_{ij} x_j + b_i \qquad (i = 1, 2, \ldots, n)$$

gegeben ist.

Ist A kontrahierend, so können wir die Methode der sukzessiven Approximation zur Lösung der Gleichung $Ax = x$ anwenden.

Unter welchen Bedingungen ist die Abbildung A nun eine Kontraktion? Die Antwort auf diese Frage hängt von der Auswahl der Metrik des Raumes ab. Wir betrachten dazu die in 2.1.1. eingeführten Räume $\boldsymbol{R}_0^n$, $\boldsymbol{R}_1^n$ und $\boldsymbol{R}^n$.

a) Im Raum $\boldsymbol{R}_0^n$ mit $\varrho(x, y) = \max\limits_{1 \leqq i \leqq n} |x_i - y_i|$ gilt

$$\begin{aligned}
\varrho(y', y'') &= \max_i |y_i' - y_i''| = \max_i \left|\sum_j a_{ij}(x_j' - x_j'')\right| \\
&\leqq \max_i \sum_j |a_{ij}|\, |x_j' - x_j''| \\
&\leqq \max_i \sum_j |a_{ij}| \max |x_j' - x_j''| \\
&= \left(\max_i \sum_j |a_{ij}|\right) \varrho(x', x'');
\end{aligned}$$

hier ergibt sich als Kontraktionsbedingung

$$\sum_{j=1}^{n} |a_{ij}| \leqq \alpha < 1, \qquad i = 1, \ldots, n. \tag{2}$$

b) Im Raum $\boldsymbol{R}_1^n$ mit $\varrho(x, y) = \sum\limits_{i=1}^{n} |x_i - y_i|$ gilt

$$\begin{aligned}
\varrho(y', y'') &= \sum_i |y_i' - y_i''| = \sum_i \left|\sum_j a_{ij}(x_j' - x_j'')\right| \\
&\leqq \sum_i \sum_j |a_{ij}|\, |x_j' - x_j''| \\
&\leqq \left(\max_j \sum_i |a_{ij}|\right) \varrho(x', x''),
\end{aligned}$$

hier lautet die Kontraktionsbedingung

$$\sum_i |a_{ij}| \leqq \alpha < 1, \qquad j = 1, \ldots, n. \tag{3}$$

c) Im Raum $\boldsymbol{R}^n$ mit $\varrho(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$ gilt auf Grund der Schwarzschen Ungleichung

$$\varrho^2(y', y'') = \sum_i \Big(\sum_j a_{ij}(x_j' - x_j'')\Big)^2 \leqq \Big(\sum_i \sum_j a_{ij}^2\Big) \varrho^2(x', x''),$$

d. h., die Kontraktionsbedingung lautet

$$\sum_i \sum_j a_{ij}^2 \leqq \alpha < 1. \tag{4}$$

Wenn also wenigstens eine der Bedingungen[1]) (2) bis (4) erfüllt ist, existiert genau ein Punkt $(x_1, x_2, \ldots, x_n)$ mit $x_i = \sum_{j=1}^{n} a_{ij}x_j + b_i$, wobei die sukzessive Approximation dieser Lösung eine Folge der Form

$$x^{(0)} = (x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)}),$$
$$x^{(1)} = (x_1^{(1)}, x_2^{(1)}, \ldots, x_n^{(1)}),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)}),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$

mit

$$x_i^{(k)} = \sum_{j=1}^{n} a_{ij}x_j^{(k-1)} + b_i$$

liefert; dabei kann man als $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$ jeden beliebigen Punkt aus dem $\boldsymbol{R}^n$ nehmen.

Jede der Bedingungen (2) bis (4) ist hinreichend dafür, daß die Abbildung $y = Ax$ eine Kontraktion ist. Für die Bedingung (2) könnte nachgewiesen werden, daß sie auch notwendig dafür ist, daß die Abbildung $y = Ax$ eine Kontraktion ist (im Sinne der Metrik a)).

Keine der Bedingungen (2) bis (4) ist jedoch notwendig für die Anwendbarkeit der Methode der sukzessiven Approximation.

---

[1]) Insbesondere folgt aus jeder der Bedingungen (2) bis (4)

$$\begin{vmatrix} a_{11} - 1 & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - 1 & \ldots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - 1 \end{vmatrix} \neq 0.$$

Wenn $|a_{ij}| < 1/n$ ist, sind alle drei Bedingungen (2) bis (4) erfüllt, und die Methode der sukzessiven Approximation ist natürlich anwendbar.

Für $|a_{ij}| \geqq 1/n$ ist keine der Bedingungen (2) bis (4) erfüllt.

2.4.3. Existenz- und Unitätssätze für Differentialgleichungen. Im vorhergehenden Abschnitt wurden zwei der einfachsten Beispiele für die Anwendung des Prinzips der kontrahierenden Abbildung im eindimensionalen und im n-dimensionalen Raum angeführt. Für die Analysis am bedeutendsten sind jedoch die Anwendungen dieses Prinzips in unendlichdimensionalen Funktionenräumen. Wir werden gleich zeigen, wie man mit seiner Hilfe Existenz- und Eindeutigkeitssätze von Lösungen einiger Typen von Differential- und Integralgleichungen erhalten kann.

1. *Cauchysche Aufgabe.* Gegeben sei die Differentialgleichung

$$\frac{dy}{dx} = f(x, y) \tag{5}$$

mit der Anfangsbedingung

$$y(x_0) = y_0. \tag{6}$$

Dabei sei die Funktion f in einem ebenen Gebiet G, das den Punkt (x_0, y_0) enthält, definiert und stetig und genüge dort bezüglich y der Lipschitzbedingung

$$|f(x, y_1) - f(x, y_2)| \leqq M\,|y_1 - y_2|.$$

Wir werden beweisen, daß dann in einem gewissen Intervall $|x - x_0| \leqq d$ eine und nur eine Lösung $y = \varphi(x)$ der Gleichung (5) existiert, die der Anfangsbedingung (6) genügt (Satz von Picard).

Die Gleichung (5) ist zusammen mit der Anfangsbedingung (6) äquivalent zu der Integralgleichung

$$\varphi(x) = y_0 + \int_{x_0}^{x} f\big(t, \varphi(t)\big)\, dt. \tag{7}$$

Wegen der Stetigkeit der Funktion f ist $|f(x, y)| \leqq K$ in einem gewissen Gebiet $G' \subset G$, das den Punkt (x_0, y_0) enthält. Wir wählen jetzt $d > 0$ so, daß die Bedingungen

1. $(x, y) \in G'$ für $|x - x_0| \leqq d$, $|y - y_0| \leqq Kd$;

2. $Md < 1$

erfüllt sind. Mit C^* bezeichnen wir dann den Raum aller stetigen Funktionen φ, die auf dem Intervall $|x - x_0| \leqq d$ definiert sind und für die $|\varphi(x) - y_0| \leqq Kd$ ist, mit dem Abstand $\varrho(\varphi_1, \varphi_2) = \max\limits_{x} |\varphi_1(x) - \varphi_2(x)|$.

Der Raum C^* ist vollständig, weil er abgeschlossener Teilraum des vollständigen Raumes aller stetigen Funktionen auf $[x_0 - d, x_0 + d]$ ist. Wir betrachten nun die Abbildung $\psi = A\varphi$, die durch die Formel

$$\psi(x) = y_0 + \int_{x_0}^{x} f\big(t, \varphi(t)\big)\, dt$$

definiert wird, wobei $|x - x_0| \leqq d$ ist. Diese Abbildung führt den vollständigen Raum C^* in sich über und erweist sich dort als Kontraktion. Ist nämlich $\varphi \in C^*$ und $|x - x_0| \leqq d$, dann ergibt sich

$$|\psi(x) - y_0| = \left| \int_{x_0}^{x} f\big(t, \varphi(t)\big)\, dt \right| \leqq Kd,$$

und daher ist $A(C^*) \subset C^*$. Außerdem gilt

$$\begin{aligned} |\psi_1(x) - \psi_2(x)| &\leqq \int_{x_0}^{x} |f\big(t, \varphi_1(t)\big) - f\big(t, \varphi_2(t)\big)|\, dt \\ &\leqq Md \max_x |\varphi_1(x) - \varphi_2(x)|, \end{aligned}$$

und wegen $Md < 1$ ist A also eine Kontraktion.

Hieraus folgt, daß die Gleichung $\varphi = A\varphi$ (d. h. Gleichung (7)) eine und nur eine Lösung im Raum C^* besitzt.

2. *Die Cauchysche Aufgabe für Gleichungssysteme.* Gegeben sei das Differentialgleichungssystem

$$\varphi_i'(x) = f_i\big(x, \varphi_1(x), \ldots, \varphi_n(x)\big), \qquad i = 1, 2, \ldots, n, \tag{8}$$

mit den Anfangsbedingungen

$$\varphi_i(x_0) = y_{0i}, \qquad i = 1, 2, \ldots, n. \tag{9}$$

Dabei seien die Funktionen f_i in einem Gebiet G des Raumes $\boldsymbol{R}^{n+1}$, das den Punkt $(x_0, y_{01}, \ldots, y_{0n})$ enthält, definiert und stetig, und außerdem mögen sie der Lipschitzbedingung

$$|f_i(x, y_1^{(1)}, \ldots, y_n^{(1)}) - f(x, y_1^{(2)}, \ldots, y_n^{(2)})| \leqq M \max_{1 \leqq i \leqq n} |y_i^{(1)} - y_i^{(2)}|$$

genügen. Wir werden zeigen, daß dann in einem gewissen Intervall $|x - x_0| \leqq d$ eine und nur eine Lösung der Anfangswertaufgabe (8), (9) existiert, d. h. ein und nur ein System von Funktionen φ_i, die den Gleichungen (8) und den Anfangsbedingungen (9) genügen.

Das System (8) ist zusammen mit den Anfangsbedingungen (9) äquivalent mit dem Integralgleichungssystem

$$\varphi_i(x) = y_{0i} + \int_{x_0}^{x} f_i\big(t, \varphi_1(t), \ldots, \varphi_n(t)\big)\, dt, \qquad i = 1, \ldots, n. \tag{10}$$

Wegen der Stetigkeit sind die Funktionen f_i in einem gewissen Gebiet $G' \subset G$, das den Punkt $(x_0, y_{01}, \ldots, y_{0n})$ enthält, beschränkt, d. h., es existiert eine positive Zahl K mit $|f_i(x, y_1, \ldots, y_n)| \leqq K$. Wir wählen jetzt $d > 0$ so, daß die folgenden Bedingungen erfüllt sind:

1. $(x, y_1, \ldots, y_n) \in G'$, wenn $|x - x_0| \leqq d$, $|y_i - y_{0i}| \leqq Kd$, $i = 1, \ldots, n$;
2. $Md < 1$.

Mit C_n^* bezeichnen wir dann den Raum aller n-Tupel $\varphi = (\varphi_1, \ldots, \varphi_n)$ von Funktionen, die für $|x - x_0| \leqq d$ definiert und stetig sind und für die

$$|\varphi_i(x) - y_{0i}| \leqq Kd$$

gilt. Den Abstand definieren wir durch die Formel

$$\varrho(\varphi, \psi) = \max_{x,i} |\varphi_i(x) - \psi_i(x)|.$$

Der so eingeführte Raum ist vollständig. Die Abbildung $\psi = A\varphi$, die durch das Gleichungssystem

$$\psi_i(x) = y_{0i} + \int_{x_0}^{x} f_i\big(t, \varphi_1(t), \ldots, \varphi_n(t)\big)\, dt$$

definiert wird, ist dann eine kontrahierende Abbildung des vollständigen Raumes C_n^* in sich. Denn aus

$$\psi_i^{(1)}(x) - \psi_i^{(2)}(x) = \int_{x_0}^{x} \Big[f_i\big(t, \varphi_1^{(1)}(t), \ldots, \varphi_n^{(1)}(t)\big) - f_i\big(t, \varphi_1^{(2)}(t), \ldots, \varphi_n^{(2)}(t)\big)\Big]\, dt$$

ergibt sich

$$\max_{x,i} |\psi_i^{(1)}(x) - \psi_i^{(2)}(x)| \leqq Md \max_{x,i} |\varphi_i^{(1)}(x) - \varphi_i^{(2)}(x)|,$$

und wegen $Md < 1$ ist A eine Kontraktion.

Hieraus folgt, daß die Operatorgleichung $\varphi = A\varphi$ genau eine Lösung im Raum C_n^* besitzt.

2.4.4. Die Anwendung des Prinzips der kontrahierenden Abbildung auf Integralgleichungen

1. *Fredholmsche Integralgleichungen.* Wir wollen die Methode der kontrahierenden Abbildung jetzt zum Beweis der Existenz und Eindeutigkeit der Lösung der in-

homogenen linearen *Fredholmschen Integralgleichung* zweiter Art verwenden. Wir betrachten also die Gleichung

$$f(x) = \lambda \int_a^b K(x, y) f(y)\, dy + \varphi(x), \tag{11}$$

in der K (der sogenannte *Kern*) und φ gegebene Funktionen, f die gesuchte Funktion und λ ein beliebiger Parameter sind.

Wir werden sehen, daß unsere Methode nur bei hinreichend kleinen Werten des Parameters λ anwendbar ist.

Zunächst setzen wir noch voraus, daß $K(x, y)$ und $\varphi(x)$ stetig sind für $a \leqq x \leqq b$, $a \leqq y \leqq b$ und somit $|K(x, y)| \leqq M$ ist. Wir betrachten nun die Abbildung $g = Af$ des vollständigen Raumes $C[a, b]$ in sich, die durch die Formel

$$g(x) = \lambda \int_a^b K(x, y) f(y)\, dy + \varphi(x)$$

gegeben ist. Dann gilt

$$\varrho(g_1, g_2) = \max |g_1(x) - g_2(x)| \leqq |\lambda|\, M(b - a) \max |f_1(x) - f_2(x)|.$$

Folglich ist A kontrahierend, sofern $|\lambda| < \dfrac{1}{M(b - a)}$ ist.

Nach dem Prinzip der kontrahierenden Abbildung können wir schließen, daß die Fredholmsche Gleichung für jedes λ mit $|\lambda| < \dfrac{1}{M(b - a)}$ eine eindeutig bestimmte stetige Lösung besitzt. Die sukzessive Approximation dieser Lösung liefert eine Folge $f_0, f_1, \ldots, f_n, \ldots$ der Form

$$f_n(x) = \lambda \int_a^b K(x, y) f_{n-1}(y)\, dy + \varphi(x),$$

wobei man für $f_0(x)$ eine beliebige stetige Funktion nehmen kann.

2. *Nichtlineare Integralgleichungen.* Das Prinzip der kontrahierenden Abbildung kann man auch auf *nichtlineare Integralgleichungen* der Form

$$f(x) = \lambda \int_a^b K\big(x, y; f(y)\big)\, dy + \varphi(x) \tag{12}$$

anwenden, wobei K und φ stetig sind und der Kern K außerdem einer Lipschitzbedingung bezüglich seines „funktionalen" Arguments genügt:

$$|K(x, y; z_1) - K(x, y; z_2)| \leqq M\, |z_1 - z_2|.$$

Für die Abbildung $g = Af$ des vollständigen Raumes $C[a, b]$ in sich, die durch die Formel

$$g(x) = \lambda \int_a^b K(x, y; f(y))\, dy + \varphi(x) \tag{13}$$

gegeben wird, gilt in diesem Fall die Ungleichung

$$\max |g_1(x) - g_2(x)| \leqq |\lambda|\, M(b - a) \max |f_1(x) - f_2(x)|,$$

wobei $g_1 = Af_1$ und $g_2 = Af_2$ ist. Folglich ist die Abbildung A für $|\lambda| < \frac{1}{M(b - a)}$ kontrahierend.

3. *Die Volterrasche Gleichung.* Wir betrachten schließlich die *Integralgleichung vom Volterraschen Typ*

$$f(x) = \lambda \int_a^x K(x, y)\, f(y)\, dy + \varphi(x). \tag{14}$$

Hier erscheint die Veränderliche x im Unterschied zur Fredholmschen Gleichung als obere Integrationsgrenze. Formal kann man diese Gleichung auch als Spezialfall der Fredholmschen Gleichung auffassen, indem man die Funktion K durch die Gleichung

$$K(x, y) = 0 \quad \text{für} \quad y > x$$

fortsetzt.

Im Fall der Fredholmschen Integralgleichung waren wir gezwungen, uns auf wenige Werte des Parameters λ zu beschränken, während das Prinzip der kontrahierenden Abbildung (und die Methode der sukzessiven Approximation) bei der Volterraschen Gleichung für alle Werte λ anwendbar ist. Genauer handelt es sich um folgende Verallgemeinerung des Prinzips der kontrahierenden Abbildung:

Es sei A eine stetige Abbildung des vollständigen metrischen Raumes R in sich, für die eine gewisse Potenz $B = A^n$ eine Kontraktion ist. Dann besitzt die Gleichung

$$Ax = x$$

genau eine Lösung.

Zum Beweis betrachten wir einen Fixpunkt x der Abbildung B, d. h. $Bx = x$. Dann haben wir

$$Ax = AB^k x = B^k Ax = B^k x_0 \to x \qquad (k \to \infty),$$

weil die Folge $Bx_0, B^2x_0, B^3x_0, \ldots$ wegen der Kontraktionseigenschaft der Abbildung B für jedes $x_0 \in R$ gegen den Fixpunkt x der Abbildung B konvergiert. Folglich ist

$$Ax = x.$$

Dieser Fixpunkt ist eindeutig bestimmt, da jeder Fixpunkt von A auch Fixpunkt der kontrahierenden Abbildung A^n ist, für die es nur einen Fixpunkt geben kann.

Wir weisen jetzt nach, daß eine gewisse Potenz der Abbildung

$$Af(x) = \lambda \int_a^x K(x, y)\, f(y)\, dy + \varphi(x)$$

eine Kontraktion ist. Dazu seien f_1 und f_2 zwei stetige Funktionen auf dem Intervall $[a, b]$. Dann gilt

$$\begin{aligned} |Af_1(x) - Af_2(x)| &= |\lambda| \left| \int_a^x K(x, y) \big(f_1(y) - f_2(y)\big)\, dy \right| \\ &\leqq |\lambda|\, M(x - a) \max |f_1(x) - f_2(x)|. \end{aligned}$$

Dabei ist

$$M = \max |K(x, y)|.$$

Hieraus folgt

$$|A^2f_1(x) - A^2f_2(x)| \leqq |\lambda|^2 M^2 \frac{(x - a)^2}{2} \max |f_1(x) - f_2(x)|$$

und allgemein

$$|A^nf_1(x) - A^nf_2(x)| \leqq |\lambda|^n M^n \frac{(x - a)^n}{n!} m \leqq |\lambda|^n M^n m \frac{(b - a)^n}{n!},$$

wobei $m = \max |f_1(x) - f_2(x)|$ ist.

Für jeden Wert von λ kann man nun die Zahl n so groß wählen, daß

$$\frac{|\lambda|^n M^n (b - a)^n}{n!} < 1$$

wird. Daher ist die Abbildung A^n eine Kontraktion. Folglich besitzt die Volterrasche Gleichung (14) für beliebiges λ eine Lösung, die darüber hinaus eindeutig bestimmt ist.

2.5. Topologische Räume

2.5.1. Definition und Beispiele topologischer Räume. Die Grundbegriffe der Theorie der metrischen Räume (Häufungspunkt, Berührungspunkt, Abschluß einer Menge usw.) haben wir eingeführt, indem wir uns auf den Begriff der Umgebung oder, was im wesentlichen dasselbe ist, auf den Begriff der offenen Menge stützten. Diese letzten beiden Begriffe (Umgebung, offene Menge) wurden ihrerseits mit Hilfe der Metrik definiert, die im betrachteten Raum gegeben war. Man kann jedoch auch einen

anderen Weg beschreiten und in einer gegebenen Menge R sofort ein System offener Mengen axiomatisch definieren, ohne in R eine Metrik einzuführen. Dieser Weg, der eine wesentlich größere Allgemeinheit garantiert, führt uns zu den *topologischen Räumen,* unter denen die metrischen Räume einen zwar sehr wichtigen, aber doch speziellen Fall darstellen.

Definition. Es sei X irgendeine Menge, die Grundmenge des Raumes. Ein beliebiges System τ von Teilmengen $G \subset X$ heißt *Topologie* in X, wenn es den beiden folgenden Forderungen genügt:

1°. Die ganze Menge und die leere Menge $\emptyset$ gehören zu τ.

2°. Die Vereinigung $\bigcup\limits_{\alpha} G_\alpha$ von beliebig vielen Mengen aus τ und der Durchschnitt $\bigcap\limits_{k=1}^{n} G_k$ von endlich vielen Mengen aus τ gehören zu τ.

Die Menge X mit einer darin gegebenen Topologie τ, d. h. ein Paar (X, τ), heißt *topologischer Raum.*

Die Mengen, die zum System τ gehören, heißen *offen.*

Ebenso wie ein metrischer Raum die Gesamtheit einer Menge von Punkten, des „Trägers", und einer darin eingeführten Metrik war, ist ein topologischer Raum die Gesamtheit einer Menge von Punkten und einer darin eingeführten Topologie. Somit bedeutet die Angabe eines topologischen Raumes die Angabe einer Menge X und die Angabe einer Topologie τ darin, d. h., es sind jene Teilmengen von X anzugeben, die als offen angesehen werden sollen.

Es ist klar, daß man in ein und derselben Menge X verschiedene Topologien einführen kann, so daß man verschiedene topologische Räume erhält. Einen topologischen Raum, d. h. ein Paar (X, τ), werden wir nur mit einem Buchstaben bezeichnen, etwa mit T. Die Elemente des topologischen Raumes nennen wir *Punkte.*

Die Mengen $T \setminus G$, die Komplemente der offenen Mengen, heißen *abgeschlossene* Mengen des topologischen Raumes T. Aus den Axiomen 1° und 2° folgt wegen der Dualitätsbeziehung (siehe 1.1.):

1. Die leere Menge $\emptyset$ und ganz T ist abgeschlossen.

2. Der Durchschnitt von beliebig (endlich oder unendlich) vielen und die Vereinigung von endlich vielen abgeschlossenen Mengen ist abgeschlossen.

Auf Grund dieser Definition ist es naheliegend, in jedem topologischen Raum die Begriffe Umgebung, Berührungspunkt, Abschluß einer Menge usw. einzuführen.

Jede offene Menge $G \subset T$, die den Punkt $x \in T$ enthält, heißt *Umgebung* von x; ein Punkt $x \in T$ heißt *Berührungspunkt* der Menge $M \subset T$, wenn jede Umgebung des Punktes x wenigstens einen Punkt aus M enthält; x heißt *Häufungspunkt* der Menge M, wenn jede Umgebung des Punktes x wenigstens einen von x verschiedenen Punkt aus M enthält. Die Gesamtheit aller Berührungspunkte der Menge M heißt *Abschluß* der Menge M und wird mit dem Symbol $[M]$ bezeichnet. Es ist leicht zu zeigen (man führe diesen Beweis), daß *die abgeschlossenen Mengen* (die von uns oben als Komplemente der offenen Mengen definiert wurden) *und nur diese die Bedingung*

$[M] = M$ erfüllen. Wie auch im Fall des metrischen Raumes ist $[M]$ die kleinste abgeschlossene Menge, die M enthält.

Aufgabe. Man zeige, daß die Abschließung, die mit Hilfe der Topologie definiert wurde, die Eigenschaften 1 bis 4 aus Satz 1, 2.2., besitzt.

Beispiele

1. Nach Satz 3′ aus 2.2. genügen die offenen Mengen in jedem metrischen Raum den Axiomen 1^0 und 2^0 der Definition eines topologischen Raumes. Somit ist jeder metrische Raum auch ein topologischer Raum.

2. Es sei T eine beliebige Menge. Als offen betrachten wir alle ihre Teilmengen. Die Axiome 1^0 und 2^0 sind dann trivialerweise erfüllt, d. h., wir erhalten einen topologischen Raum. Darin sind alle Mengen gleichzeitig offen und abgeschlossen, und das bedeutet, daß jede Menge mit ihrem Abschluß zusammenfällt. Eine solche triviale Topologie besitzt zum Beispiel der in 2.1., Beispiel 1, angegebene metrische Raum.

3. Als anderen Extremfall betrachten wir in einer beliebigen Menge X die Topologie, die nur aus zwei Mengen besteht, aus ganz X und der leeren Menge $\emptyset$. Hier ist der Abschluß jeder nichtleeren Menge ganz X. Einen solchen topologischen Raum kann man als „Raum *zusammengeklebter* Punkte" bezeichnen.

4. T möge jetzt aus zwei Punkten a und b bestehen. Zu den offenen Mengen zählen wir diesmal T, die leere Menge und die Menge, die nur aus dem einen Punkt b besteht. Die Axiome 1^0 und 2^0 sind hier erfüllt. In diesem Raum (den man oft als *zusammenhängendes Punktepaar* bezeichnet) sind folgende Mengen abgeschlossen: ganz T, die leere Menge und der Punkt a. Der Abschluß der einpunktigen Menge $\{b\}$ ist ganz T.

Aufgabe. Man konstruiere alle Topologien in einem Raum X, der aus zwei, drei, vier und fünf Punkten besteht.

2.5.2. Der Vergleich von Topologien. In ein und derselben Trägermenge X seien zwei Topologien τ_1 und τ_2 gegeben (dadurch werden zwei topologische Räume definiert: $T_1 = (X, \tau_1)$ und $T_2 = (X, \tau_2)$). Wir sagen, daß die Topologie τ_1 *stärker* oder *feiner* als die Topologie τ_2 ist, wenn das Mengensystem τ_2 in τ_1 enthalten ist. Von der Topologie τ_2 sagt man dabei, daß sie *schwächer* oder *gröber* als τ_1 ist.

In der Gesamtheit aller möglichen Topologien einer Menge X kann damit auf natürliche Weise eine Halbordnung eingeführt werden (die Topologie τ_2 geht der Topologie τ_1 voraus, wenn sie schwächer ist als τ_1). In dieser Gesamtheit von Topologien gibt es ein maximales Element, nämlich die Topologie, in der alle Mengen offen sind (Beispiel 2), und ein minimales Element, nämlich die Topologie, in der nur ganz X und $\emptyset$ offen sind (Beispiel 3).

Satz 1. *Der Durchschnitt einer beliebigen Menge von Topologien $\tau = \bigcap_\alpha \tau_\alpha$ in X ist ebenfalls eine Topologie in X. Diese Topologie τ ist schwächer als jede Topologie τ_α.*

Beweis. Es ist klar, daß $\bigcap_\alpha \tau_\alpha$ die Mengen X und $\emptyset$ enthält. Daraus, daß jedes τ_α bezüglich der Bildung beliebiger Vereinigungen und endlicher Durchschnitte abgeschlossen ist, folgt ferner, daß auch $\tau = \bigcap_\alpha \tau_\alpha$ diese Eigenschaft besitzt.

Folgerung. *Es sei $\mathfrak{B}$ eine beliebige Menge von Teilmengen der Menge X; dann existiert eine minimale Topologie in X, die $\mathfrak{B}$ enthält.*

Es gibt nämlich Topologien, die $\mathfrak{B}$ enthalten (z. B. jene, in der alle $A \subset X$ offen sind), und der Durchschnitt aller Topologien, die $\mathfrak{B}$ enthalten, ist die gesuchte Topologie. Diese minimale Topologie heißt die *durch das System* $\mathfrak{B}$ *erzeugte* Topologie und wird mit $\tau(\mathfrak{B})$ bezeichnet.

Es sei X eine beliebige Menge und A eine Teilmenge von X. Dann heißt das System $\mathfrak{B}_A$, das aus allen Teilmengen der Form $A \cap B$, $B \in \mathfrak{B}$, besteht, die *Spur* des Mengensystems $\mathfrak{B}$ auf der Teilmenge A. Es ist leicht zu sehen, daß (auf A) die Spur der Topologie τ (die in X gegeben ist) eine Topologie τ_A in A ist. Somit erweist sich jede Teilmenge A eines beliebigen topologischen Raumes selbst wieder als topologischer Raum. Der topologische Raum (A, τ_A) heißt *Teilraum* des ursprünglichen topologischen Raumes (X, τ). Es ist klar, daß zwei verschiedene Topologien τ_1 und τ_2 in X in $A \subset X$ ein und dieselbe Topologie erzeugen können. Die Topologie τ_A heißt die in A *induzierte Topologie.*

2.5.3. Erzeugende Umgebungssysteme. Basis. Abzählbarkeitsaxiome. Eine Topologie in einem Raum T anzugeben heißt, wie wir sahen, in diesem Raum ein System offener Mengen auszuzeichnen. Jedoch ist es in konkreten Aufgaben bequemer, nicht die ganze Topologie anzugeben, sondern nur einen gewissen Teil davon, d. h. eine gewisse Teilmenge der offenen Mengen, durch die die Gesamtheit aller offenen Mengen eindeutig bestimmt ist. So haben wir zum Beispiel im metrischen Raum zunächst den Begriff der offenen Kugel (ε-Umgebung) eingeführt und anschließend die offenen Mengen als jene Mengen definiert, in denen jeder Punkt zusammen mit einer gewissen Kugelumgebung enthalten ist. Mit anderen Worten sind im metrischen Raum genau die Mengen offen, die man als Vereinigung von (endlich oder unendlich vielen) Kugeln darstellen kann. Insbesondere sind auf der Zahlengeraden genau die Mengen offen, die sich als Vereinigung von offenen Intervallen darstellen lassen. Diese Überlegungen führen uns zum wichtigen Begriff der Basis eines topologischen Raumes.

Definition. Ein System $\mathcal{B}$ von offenen Mengen heißt *Basis* eines topologischen Raumes T, wenn jede offene Menge in T als Vereinigung von (endlich oder unendlich vielen) Mengen aus $\mathcal{B}$ dargestellt werden kann.

So bildet zum Beispiel die Gesamtheit aller offenen Kugeln (mit beliebigen Mittelpunkten und Radien) eine Basis im metrischen Raum. Insbesondere ist das System aller offenen Intervalle eine Basis auf der Zahlengeraden. Eine Basis auf der Zahlengeraden bilden auch schon die offenen Intervalle mit rationalen Endpunkten allein,

da man ein beliebiges offenes Intervall und damit auch eine beliebige offene Menge auf der Zahlengeraden als Vereinigung von solchen Intervallen darstellen kann.

Somit kann man die Topologie τ im Raum T auch einführen, indem man in diesem Raum eine Basis $\mathcal{G}$ angibt; die Topologie τ stimmt dann mit der Gesamtheit der Mengen überein, die als Vereinigung von Mengen aus $\mathcal{G}$ darstellbar sind.

Jede Basis $\mathcal{G}$ in einem topologischen Raum $T = (X, \tau)$ besitzt die folgenden beiden Eigenschaften:

1. *Jeder Punkt $x \in X$ ist in wenigstens einem $G \in \mathcal{G}$ enthalten.*

2. *Wenn x im Durchschnitt zweier Mengen G_1 und G_2 aus $\mathcal{G}$ enthalten ist, dann existiert eine Menge G_3 aus $\mathcal{G}$ mit*

$$x \in G_3 \subset G_1 \cap G_2.$$

Eigenschaft 1 bedeutet einfach, daß die offene Menge X als Vereinigung von Mengen aus $\mathcal{G}$ dargestellt werden kann, und Eigenschaft 2 ergibt sich daraus, daß $G_1 \cap G_2$ offen ist und folglich die Vereinigung gewisser Elemente der Basis ist.

Umgekehrt sei X jetzt eine beliebige Menge und $\mathcal{G}$ ein System von Teilmengen von X, das die Eigenschaften 1 und 2 besitzt. Dann bildet die Gesamtheit der Mengen, die als Vereinigung von Mengen aus $\mathcal{G}$ darstellbar sind, in X eine Topologie (d. h., sie genügt den Axiomen 1^0 und 2^0 der Definition eines topologischen Raumes).

Wenn nämlich $\tau(\mathcal{G})$ die Gesamtheit aller der Teilmengen von X ist, die sich als Vereinigung von Mengen aus $\mathcal{G}$ darstellen lassen, dann gehören die leere Menge[1]) und ganz X zu $\tau(\mathcal{G})$, und eine beliebige Vereinigung von Mengen aus $\tau(\mathcal{G})$ gehört ebenfalls zu $\tau(\mathcal{G})$. Wir werden zeigen, daß auch der Durchschnitt endlich vieler Mengen aus $\tau(\mathcal{G})$ zu $\tau(\mathcal{G})$ gehört. Dazu ist es ausreichend, das für zwei Mengen nachzuprüfen. Es sei also $A = \bigcup_{\alpha} G_\alpha$ und $B = \bigcup_{\beta} G_\beta$, dann ist $A \cap B = \bigcup_{\alpha,\beta} (G_\alpha \cap G_\beta)$. Aus Bedingung 2 folgt, daß jeder Durchschnitt $G_\alpha \cap G_\beta$ in $\tau(\mathcal{G})$ enthalten ist. Dann ist aber auch $A \cap B \in \tau(\mathcal{G})$.

Somit erhalten wir folgendes Resultat.

Satz 2. *Ein System $\mathcal{G}$ von Teilmengen G der Menge X ist genau dann Basis einer Topologie in X, wenn $\mathcal{G}$ die Eigenschaften* 1 *und* 2 *besitzt.*

Es sei jetzt im Raum T eine feste Topologie τ gegeben. Wenn wir nun in T ein System offener Mengen mit den Eigenschaften 1 und 2 als Basis nehmen, erhalten wir offensichtlich in T eine Topologie $\tau(\mathcal{G})$, die entweder mit der Ausgangstopologie τ übereinstimmt oder schwächer ist. Wir wollen Bedingungen aufstellen, unter denen $\mathcal{G}$ gerade die gegebene Topologie τ erzeugt.

Satz 3. *Ein System $\mathcal{G} \subset \tau$ ist genau dann Basis einer gegebenen Topologie τ, wenn folgende Bedingung gilt:*

3. *Für jede offene Menge G und jeden Punkt $x \in G$ existiert eine Menge $G_x \in \mathcal{G}$ mit $x \in G_x \subset G$.*

---

[1]) Man erhält sie als Vereinigung der leeren Menge von Elementen des Systems $\mathcal{G}$.

Beweis. Wenn Bedingung 3 erfüllt ist, kann jede offene Menge G in der Form

$$G = \bigcup_{x \in G} G_x$$

dargestellt werden, d. h., $\mathscr{G}$ ist Basis der Topologie τ. Wenn umgekehrt $\mathscr{G}$ Basis der Topologie τ ist, kann jedes $G \in \tau$ als Vereinigung von Mengen aus $\mathscr{G}$ dargestellt werden, und dann findet sich zu jedem $x \in G$ eine Menge $G_x \in \mathscr{G}$ mit $x \in G_x \subset G$.

Aufgabe. Es seien $\mathscr{G}_1$ und $\mathscr{G}_2$ zwei Basen in X (d. h. zwei Mengensysteme, die den Bedingungen 1 und 2 von S. 91 genügen) und τ_1 und τ_2 die durch sie definierten Topologien. Man zeige, daß $\tau_1 \subset \tau_2$ dann und nur dann gilt, wenn für jedes $G_1 \in \mathscr{G}$ und jeden Punkt $x \in G_1$ ein $G_2 \in \mathscr{G}_2$ existiert, so daß $x \in G_2 \subset G_1$ ist.

Mit Hilfe von Satz 3 kann man zum Beispiel leicht feststellen, daß die Gesamtheit aller offenen Kugeln in einem metrischen Raum eine Basis seiner Topologie bildet. Die Gesamtheit aller Kugeln mit rationalen Radien stellt ebenfalls eine Basis dar. Auf der Zahlengeraden ist etwa die Gesamtheit aller rationalen Intervalle (d. h. Intervalle mit rationalen Endpunkten) Beispiel einer Basis.

Eine wichtige Klasse von topologischen Räumen bilden die *Räume mit abzählbarer Basis*, d. h. Räume, in denen wenigstens eine Basis existiert, die aus höchstens abzählbar vielen Mengen besteht. Besitzt ein topologischer Raum eine abzählbare Basis, so sagt man auch, daß er dem *zweiten Abzählbarkeitsaxiom* genügt.

Wenn es in einem topologischen Raum T eine abzählbare Basis gibt, dann gibt es in T auch notwendig eine abzählbare überall dichte Menge, d. h. eine abzählbare Menge, deren Abschluß ganz T ist. Wenn nämlich $\{G_n\}$ eine solche Basis ist, können wir aus jedem Element dieser Basis einen beliebigen Punkt x_n auswählen. Und die abzählbare Menge $X = \{x_n\}$ ist dann überall dicht in T, weil andernfalls die nichtleere offene Menge $G = T \setminus [X]$ keinen Punkt aus X enthalten würde. Das ist aber nicht möglich, da G die Vereinigung gewisser Mengen des Systems $\{G_n\}$ ist und $x_n \in G_n$.

Topologische wie auch metrische Räume mit einer abzählbaren überall dichten Menge heißen *separabel*.

Für metrische Räume ist auch die Umkehrung des eben Gezeigten richtig:

Wenn ein metrischer Raum R separabel ist, dann gibt es in R eine abzählbare Basis. Eine solche Basis bilden zum Beispiel die offenen Kugeln $B(x_n, 1/m)$, wobei $\{x_n\}$ die abzählbare überall dichte Menge ist und n und m unabhängig voneinander alle natürlichen Zahlen durchlaufen. Damit gilt folgender Satz:

Satz 4. *Ein metrischer Raum R besitzt genau dann eine abzählbare Basis, wenn er separabel ist.*

Nach diesem Satz können alle Beispiele separabler metrischer Räume auch als Beispiele von metrischen Räumen mit abzählbarer Basis dienen, während der nichtseparable Raum der beschränkten Folgen (siehe 2.1., Beispiel 9) keine abzählbare Basis besitzt.

Bemerkung. Satz 4 ist nicht allgemein für beliebige (nicht metrische) topologische Räume richtig. Man kann Beispiele separabler Räume ohne abzählbare Basis angeben. Wir wollen die hier auftretenden Erscheinungen klären. Für jeden Punkt x des metrischen Raumes R existiert ein abzählbares System $\mathfrak{U}$ von Umgebungen von x (zum Beispiel das System der offenen Kugeln $B(x, 1/n)$), das folgende Eigenschaft besitzt: Zu jeder offenen Menge G, die den Punkt x enthält, findet sich eine Umgebung aus dem System $\mathfrak{U}$, die vollständig in G liegt. Ein solches Umgebungssystem heißt *Umgebungsbasis* des Punktes x.

Wenn ein Punkt x des topologischen Raumes T eine abzählbare Umgebungsbasis besitzt, sagt man, daß in diesem Punkt das *erste Abzählbarkeitsaxiom* erfüllt ist. Wenn das für jeden Punkt des Raumes T gültig ist, sagt man, daß T dem ersten Abzählbarkeitsaxiom genügt.

Jeder metrische Raum genügt automatisch dem ersten Abzählbarkeitsaxiom, selbst wenn er nicht separabel ist. Jedoch in einem beliebigen topologischen Raum (sogar wenn er nur aus abzählbar vielen Punkten besteht) braucht das erste Abzählbarkeitsaxiom nicht zu gelten. Deswegen übertragen sich jene Überlegungen, mit deren Hilfe wir für einen metrischen Raum aus dem Vorhandensein einer abzählbaren überall dichten Menge die Existenz einer abzählbaren Basis in einem solchen Raum folgerten, nicht auf den Fall eines beliebigen topologischen Raumes. Sogar in einem separablen topologischen Raum, der dem ersten Abzählbarkeitsaxiom genügt, braucht keine abzählbare Basis zu existieren.

Ein Mengensystem $\{M_\alpha\}$ heißt *Überdeckung* der Menge X, wenn $\bigcup_\alpha M_\alpha = X$ ist. Eine Überdeckung des topologischen Raumes T, die aus offenen (abgeschlossenen) Mengen besteht, heißt *offene* (*abgeschlossene*) Überdeckung. Wenn ein Teilsystem $\{M_{\alpha_i}\}$ der Überdeckung $\{M_\alpha\}$ bereits eine Überdeckung des Raumes T bildet, dann heißt $\{M_{\alpha_i}\}$ *Teilüberdeckung* der Überdeckung $\{M_\alpha\}$.

Satz 5. *Ist T ein topologischer Raum mit abzählbarer Basis, dann kann man aus jeder offenen Überdeckung von T eine endliche oder abzählbare Teilüberdeckung auswählen.*

Beweis. Es sei $\{O_\alpha\}$ eine Überdeckung des Raumes T. Also ist jeder Punkt $x \in T$ in einer gewissen Umgebung O_α enthalten. $\{G_n\}$ sei die abzählbare Basis in T. Für jedes $x \in T$ existiert dann ein Element $G_n(x)$ dieser Basis mit $x \in G_n(x) \subset O_\alpha$. Die Gesamtheit der auf diese Weise ausgewählten Mengen $G_n(x)$ ist endlich oder abzählbar und überdeckt ganz T. Wenn wir nun zu jedem $G_n(x)$ eine Menge $O_\alpha \supset G_n(x)$ auswählen, erhalten wir eine endliche oder abzählbare Überdeckung von T. Damit ist der Satz bewiesen.

Nach Definition eines topologischen Raumes sind die leere Menge und der ganze Raum T gleichzeitig offen und abgeschlossen. Ein Raum, in dem keine weiteren Mengen gleichzeitig offen und abgeschlossen sind, heißt *zusammenhängend*. Die Zahlgerade $\boldsymbol{R}^1$ stellt eines der einfachsten Beispiele eines zusammenhängenden

Raumes dar. Wenn man jedoch aus dem R^1 einen Punkt entfernt, ist der verbleibende Raum nicht mehr zusammenhängend.

2.5.4. Konvergente Folgen in T. Der im Fall eines metrischen Raumes bekannte Begriff der konvergenten Folge läßt sich leicht auf topologische Räume übertragen. Und zwar heißt eine Folge $x_1, x_2, \ldots, x_n, \ldots$ von Punkten aus T *konvergent gegen den Punkt* x, wenn jede Umgebung des Punktes x alle Punkte dieser Folge von einem gewissen Index an enthält. In topologischen Räumen spielt der Begriff der Konvergenz jedoch nicht jene fundamentale Rolle, die er in metrischen Räumen besitzt. Das liegt daran, daß ein Punkt x in einem metrischen Raum R dann und nur dann Berührungspunkt einer Menge $M \subset R$ ist, wenn in M eine Folge existiert, die gegen x konvergiert, während das in einem topologischen Raum im allgemeinen nicht so ist. Daraus, daß x Berührungspunkt von M im topologischen Raum T ist (d. h. zu $[M]$ gehört), folgt nicht die Existenz einer Folge in M, die gegen x konvergiert. Als Beispiel nehmen wir etwa das abgeschlossene Intervall $[0, 1]$ und nennen jene Teilmengen (zusammen mit der leeren Menge) offen, die man aus $[0, 1]$ durch Entnahme von endlich oder abzählbar vielen Punkten erhält. Es ist leicht nachzuprüfen, daß hier die Forderungen 1^0 und 2^0 von S. 88 erfüllt sind, d. h., wir erhalten einen topologischen Raum. In diesem Raum sind nur die stationären Folgen konvergent, d. h. Folgen, deren Folgenglieder von einem gewissen Index an übereinstimmen: $x_n = x_{n+1} = \ldots$ (man zeige das!). Ist M nun beispielsweise das halboffene Intervall $[0, 1]$, dann ist der Punkt 0 zwar Häufungspunkt von M (man prüfe das nach!), aber keine Folge von Punkten aus M konvergiert bezüglich der betrachteten Topologie gegen 0.

Die konvergenten Folgen behalten ihre Bedeutung, wenn wir nicht beliebige topologische Räume betrachten, sondern nur Räume, die dem ersten Abzählbarkeitsaxiom genügen, d. h., wenn zu jedem Punkt x des Raumes T ein abzählbares erzeugendes Umgebungssystem existiert. In diesem Fall kann jeder Berührungspunkt x einer beliebigen Menge $M \subset T$ als Grenzwert einer gewissen Folge von Punkten aus M dargestellt werden. Um das zu zeigen, sei $\{O_n\}$ ein abzählbares erzeugendes Umgebungssystem des Punktes x. Dabei kann man annehmen, daß $O_{n+1} \subset O_n$ ist (sonst ersetzen wir O_n durch $\bigcap_{k=1}^{n} O_k$). In O_k $(k = 1, 2, \ldots)$ wählen wir nun einen Punkt $x_k \in M$. Es ist klar, daß ein solches x_k existiert, sonst wäre x nicht Berührungspunkt von M. Die Folge $\{x_k\}$ konvergiert aber offensichtlich gegen x.

Alle metrischen Räume genügen, wie wir festgestellt haben, dem ersten Abzählbarkeitsaxiom. Gerade deshalb konnten wir auch solche Begriffe wie Abschluß, Berührungspunkt usw. für metrische Räume mit Hilfe des Konvergenzbegriffes von Folgen formulieren.

2.5.5. Stetige Abbildungen. Homöomorphismen. Der in 2.1. für metrische Räume eingeführte Begriff der stetigen Abbildung läßt sich in natürlicher Weise auf beliebige topologische Räume verallgemeinern.

Definition. X und Y seien zwei topologische Räume. Eine Abbildung f des Raumes X in den Raum Y heißt *stetig im Punkt* x_0, wenn es für jede Umgebung U_{y_0} des Punktes $y_0 = f(x_0)$ eine Umgebung V_{x_0} gibt mit $f(V_{x_0}) \subset U_{y_0}$. Die Abbildung $f: X \to Y$ heißt *stetig*, wenn sie in jedem Punkt $x \in X$ stetig ist. Insbesondere nennen wir eine stetige Abbildung des topologischen Raumes X in die reellen Zahlen *stetige Funktion* auf diesem Raum.

Man überzeugt sich leicht davon, daß diese Definition für metrische Räume tatsächlich mit jener Stetigkeitsdefinition für Abbildungen eines metrischen Raumes in einen anderen übereinstimmt, die in 2.1. angegeben wurde.

Die obige Definition besitzt lokalen Charakter, die Stetigkeit einer Abbildung f auf dem ganzen Raum X wird durch die Stetigkeit von f in jedem Punkt definiert. Es zeigt sich, daß man den Begriff der Stetigkeit einer Abbildung eines topologischen Raumes in einen anderen auch mit Hilfe der offenen Mengen formulieren kann, d. h. mit Hilfe der Topologie dieses Raumes.

Satz 6. *Eine Abbildung f des topologischen Raumes X in den topologischen Raum Y ist genau dann stetig, wenn das Urbild $\Gamma = f^{-1}(G)$ jeder offenen Menge $G \subset Y$ (in X) offen ist.*

Beweis. Notwendigkeit. Die Abbildung f sei stetig, und G sei eine offene Menge in Y. Wir wollen beweisen, daß $\Gamma = f^{-1}(G)$ offen ist. Dazu sei x ein beliebiger Punkt der Menge Γ und $y = f(x)$. Dann ist G eine Umgebung des Punktes y. Nach Definition der Stetigkeit gibt es daher eine Umgebung V_x des Punktes x mit $f(V_x) \subset G$, d. h. $V_x \subset \Gamma$. Anders ausgedrückt, existiert zu jedem $x \in \Gamma$ eine Umgebung V_x dieses Punktes, die in Γ enthalten ist. Das bedeutet aber gerade, daß Γ offen ist.

Hinlänglichkeit. Es sei $\Gamma = f^{-1}(G)$ offen, wenn $G \subset Y$ offen ist. Wir betrachten einen beliebigen Punkt $x \in X$ und eine beliebige Umgebung U_y des Punktes $y = f(x)$. Wegen $y \in U_y$ gehört der Punkt x zur Menge $f^{-1}(U_y)$. Diese offene Menge ist aber eine Umgebung des Punktes x, deren Bild ganz in U_y enthalten ist. Damit ist der Satz bewiesen.

Bemerkung. X und Y seien beliebige Mengen und f eine Abbildung von X in Y. Wenn in Y eine Topologie gegeben ist (d. h. ein Mengensystem, das $\emptyset$ und Y enthält und bezüglich der Bildung beliebiger Vereinigungen und endlicher Durchschnitte abgeschlossen ist), dann ist das Urbild der Topologie τ (d. h. die Gesamtheit aller Mengen $f^{-1}(G)$, $G \in \tau$) eine Topologie in X.

Zum Beweis genügt es, an den Satz über das Urbild der Vereinigung und des Durchschnitts von Mengen (siehe 1.2.) zu erinnern. Wir bezeichnen diese Topologie mit $f^{-1}(\tau)$. Wenn jetzt X und Y topologische Räume mit den Topologien τ_x und τ_y sind, dann kann man Satz 6 auch so formulieren: *Die Abbildung $f: X \to Y$ ist dann und nur dann stetig, wenn die Topologie τ_x stärker als die Topologie $f^{-1}(\tau_y)$ ist.*

Weil das Urbild des Komplements gleich dem Komplement des Urbilds ist, ergibt sich ein zu Satz 6 dualer Satz.

Satz 6′. *Eine Abbildung f des topologischen Raumes X in den topologischen Raum Y ist genau dann stetig, wenn das Urbild jeder abgeschlossene Menge aus Y (in X) abgeschlossen ist.*

Man überzeugt sich leicht, daß das Bild einer offenen (abgeschlossenen) Menge bei einer stetigen Abbildung nicht notwendig offen (abgeschlossen) ist. Wir betrachten zum Beispiel die Abbildung des halboffenen Intervalls $[0, 1)$ auf die Kreislinie. Die Menge $[1/2, 1)$, die in $[0, 1)$ abgeschlossen ist, geht dabei in eine nichtabgeschlossene Menge auf der Kreislinie über (Abb. 11).

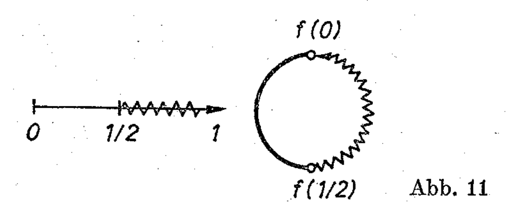


Abb. 11

Für stetige Abbildungen gilt folgender Satz, der dem aus der Analysis wohlbekannten Satz über die Stetigkeit einer zusammengesetzten Funktion analog ist.

Satz 7. *Es seien X, Y und Z topologische Räume und f und φ stetige Abbildungen von X in Y bzw. von Y in Z. Dann ist die Abbildung $x \to \varphi(f(x))$ des Raumes X in Z stetig.*

Der Beweis dieses Satzes ergibt sich sofort aus Satz 6.

Den Begriff des Homöomorphismus, der in 2.1. für metrische Räume eingeführt wurde, kann auch auf topologische Räume ausgedehnt werden. Und zwar heißt eine Abbildung f des topologischen Raumes X auf den topologischen Raum Y *Homöomorphismus,* wenn sie eineindeutig ist und wenn f und f^{-1} stetig sind; die Räume X und Y werden dabei *homöomorph* genannt. Homöomorphe Räume besitzen ein und dieselben topologischen Eigenschaften, und vom topologischen Standpunkt kann man sie einfach als zwei Exemplare ein und desselben Raumes betrachten. Die Topologien zweier homöomorpher Räume sind dabei Bild und Urbild voneinander. Die Beziehung der Homöomorphie ist reflexiv, symmetrisch und transitiv, deshalb zerfällt die Gesamtheit aller topologischen Räume in disjunkte Klassen zueinander homöomorpher Räume.

Bemerkung. Man muß jedoch berücksichtigen, daß die metrischen Eigenschaften zweier zueinander homöomorpher metrischer Räume verschieden sein können.[1]) So kann etwa einer von ihnen vollständig sein, der andere aber nicht. Zum Beispiel ist das offene Intervall $(-\pi/2, \pi/2)$ homöomorph zur Zahlengeraden (einen ent-

---

[1]) Die Metrik des Raumes R bestimmt eindeutig die Topologie in R, aber nicht umgekehrt: Die Topologie in einem metrischen Raum (R, ϱ) kann man durch Angabe verschiedener Metriken in X erhalten.

sprechenden Homöomorphismus kann man mittels der Funktion $x \to \tan x$ angeben), aber dabei ist die Gerade ein vollständiger Raum und das offene Intervall nicht.

2.5.6. Trennungsaxiome. Obwohl sich viele Grundbegriffe aus der Theorie der metrischen Räume leicht auf beliebige topologische Räume übertragen lassen, sind diese Räume ohne irgendwelche Zusatzeigenschaften für die Belange der Analysis zu allgemein. In solchen Räumen können Fälle auftreten, die sich wesentlich von den in metrischen Räumen möglichen Fällen unterscheiden. So sahen wir, daß eine endliche Menge von Punkten in einem topologischen Raum nicht abgeschlossen zu sein braucht (Beispiel 4, S. 89) usw.

Unter den topologischen Räumen kann man Räume auswählen, die den metrischen Räumen in ihren Eigenschaften näherstehen. Dazu muß man außer den Axiomen 1^0 und 2^0 eines topologischen Raumes (S. 88) noch gewisse Zusatzbedingungen verlangen. Solche Bedingungen sind zum Beispiel die Abzählbarkeitsaxiome; sie gestatten, die Topologie eines Raumes auf der Grundlage des Konvergenzbegriffes zu untersuchen. Einen weiteren wichtigen Typ von Zusatzbedingungen bilden Forderungen anderer Natur, die sogenannten *Trennungsaxiome*. Wir werden diese Axiome nach dem Grad ihrer schrittweisen Verschärfung aufzählen.

T_1-Axiom (*Erstes Trennungsaxiom*): Für je zwei verschiedene Punkte x und y des Raumes T existieren eine Umgebung O_x des Punktes x, die den Punkt y nicht enthält, und eine Umgebung O_y des Punktes y, die den Punkt x nicht enthält.

Räume, die diesem Axiom genügen, heißen *T_1-Räume*. Als Beispiel eines topologischen Raumes, der kein T_1-Raum ist, kann das zusammenhängende Punktepaar dienen.

In einem T_1-Raum ist jeder Punkt eine abgeschlossene Menge. Denn wenn $x \neq y$ ist, so existiert eine Umgebung O_y des Punktes y, die x nicht enthält, d. h., es ist $y \notin [x]$. Daher gilt $[x] = x$. Folglich ist in einem T_1-Raum auch jede endliche Punktmenge abgeschlossen. Darüber hinaus ist das T_1-Axiom sogar gleichwertig mit der Forderung nach der Abgeschlossenheit aller dieser Mengen, wie man leicht nachprüfen kann.

Weiter oben (S. 88) haben wir einen Häufungspunkt x einer Menge M im topologischen Raum T als einen Punkt definiert, für den die Menge $U \cap M \setminus \{x\}$ nicht leer ist, wenn U eine beliebige Umgebung des Punktes x ist.

In Räumen, die dem T_1-Axiom nicht genügen, können sogar endliche Mengen M einen Häufungspunkt besitzen. Wenn zum Beispiel T das zusammenhängende Punktepaar mit der Topologie ist, die aus $\emptyset$, $\{b\}$, $\{a, b\}$ besteht, dann erweist sich der Punkt a als Häufungspunkt der Menge $\{b\}$.

In T_1-Räumen kann diese Erscheinung schon nicht mehr auftreten. Es gilt nämlich folgende Beziehung.

Lemma. *Ein Punkt x ist genau dann Häufungspunkt der Menge M in einem T_1-Raum, wenn jede Umgebung U dieses Punktes unendlich viele Punkte aus M enthält.*

Die Hinlänglichkeit dieser Bedingung ist offensichtlich. Wir wollen die Notwendigkeit nachweisen. Dazu sei x ein Häufungspunkt der Menge M, und wir nehmen an, daß eine Umgebung U des Punktes x existiert, die nur endlich viele Punkte aus M enthält. Es seien $x_1, x_2, \ldots, x_n$ alle diese Punkte ohne den Punkt x (wenn dieser zu M gehört). Dann ist $V = U \setminus \{x_1, \ldots, x_n\}$ eine Umgebung von x mit $V \cap M \setminus \{x\} = \emptyset$.

Jeder metrische Raum ist offenbar ein T_1-Raum. Deshalb wurde zur Definition des Häufungspunktes einer Menge im metrischen Raum gerade die Eigenschaft verwendet, die im Lemma gezeigt wurde.

Eine Verschärfung des ersten Trennungsaxioms ist das T_2-Axiom.

T_2-Axiom (Zweites oder Hausdorffsches Trennungsaxiom): Je zwei verschiedene Punkte x und y des topologischen Raumes T besitzen disjunkte Umgebungen O_x und O_y.

Ein Raum, der diesem Axiom genügt, heißt *T_2-Raum* oder *Hausdorffraum*. Jeder Hausdorffraum ist T_1-Raum, aber nicht umgekehrt. Als Beispiel eines nicht Hausdorffschen T_1-Raumes kann das abgeschlossene Intervall $[0, 1]$ dienen, in dem die leere Menge und alle die Mengen offen sein sollen, die aus $[0, 1]$ durch Herausnahme höchstens abzählbar vieler Punkte entstehen.

T_3-Axiom (Drittes Trennungsaxiom): Ein beliebiger Punkt und jede abgeschlossene Menge, die diesen Punkt nicht enthält, besitzen disjunkte Umgebungen. *Umgebung einer Menge M* in einem topologischen Raum T heißt dabei jede offene Menge U, die M enthält.

Diesem Axiom kann man folgende äquivalente Formulierung geben:

Jede Umgebung U eines beliebigen Punktes x enthält eine kleinere Umgebung dieses Punktes, die zusammen mit ihrem Abschluß ganz in U liegt. Der Leser möge das als Übung beweisen.

Da in einem beliebigen topologischen Raum ein Punkt keine abgeschlossene Menge zu sein braucht, ist das dritte Trennungsaxiom nur für die Räume interessant, die auch dem T_1-Axiom genügen. Räume, die sowohl dem ersten als auch dem dritten Trennungsaxiom genügen, heißen *regulär*.

Jeder reguläre Raum ist offenbar Hausdorffraum. Als Beispiel eines nicht regulären Hausdorffraumes kann wieder das abgeschlossene Intervall $[0, 1]$ dienen, wenn die Topologie darin folgendermaßen definiert wird. Die Umgebungen aller von 0 verschiedenen Punkte sollen in der üblichen Weise erklärt sein, während Umgebungen des Nullpunktes alle halboffenen Intervalle $[0, \alpha)$ sein sollen, aus denen sämtliche Punkte der Form $1/n$ $(n = 1, 2, \ldots)$ entfernt wurden. In diesem Hausdorffraum sind der Punkt 0 und die Folge $\{1/n\}$ disjunkte abgeschlossene Mengen, aber sie lassen sich nicht durch disjunkte Umgebungen voneinander trennen.

Gewöhnlich sind reguläre Räume die allgemeinsten Räume, denen man in der Analysis begegnet. Vom Standpunkt der Analysis sind in der Regel sogar nur die sogenannten normalen Räume von Interesse, die die folgende noch stärkere Forderung erfüllen.

T_4-Axiom (Normalität): Ein T_1-Raum heißt *normal*, wenn darin je zwei disjunkte abgeschlossene Mengen disjunkte Umgebungen besitzen.

Zu den normalen Räumen gehören insbesondere alle metrischen Räume. Um das zu zeigen, seien X und Y zwei disjunkte abgeschlossene Mengen im metrischen Raum R. Jeder Punkt $x \in X$ besitzt eine Umgebung O_x, die disjunkt zu Y ist, und folglich befindet sich x in einem gewissen positiven Abstand ϱ_x von Y. Analog ist der Abstand jedes Punktes $y \in Y$ von X eine positive Größe ϱ_y. Wir betrachten nun die offenen Mengen[1])

$$U = \bigcup_{x \in X} B\left(x, \frac{\varrho_x}{2}\right) \quad \text{und} \quad V = \bigcup_{y \in Y} B\left(y, \frac{\varrho_y}{2}\right),$$

die X bzw. Y enthalten, und werden zeigen, daß ihr Durchschnitt leer ist. Dazu nehmen wir an, daß $z \in U \cap V$ ist. Dann existiert in X ein Punkt x_0 mit $\varrho(x_0, z) < \varrho_{x_0}/2$ und in Y ein Punkt y_0 mit $\varrho(z, y_0) < \varrho_{y_0}/2$. Dabei sei etwa $\varrho_{x_0} \leqq \varrho_{y_0}$. Dann ist

$$\varrho(x_0, y_0) \leqq \varrho(x_0, z) + \varrho(z, y_0) < \frac{\varrho_{x_0}}{2} + \frac{\varrho_{y_0}}{2} \leqq \varrho_{y_0},$$

d. h. $x_0 \in B(y_0, \varrho_{y_0})$. Das widerspricht aber der Definition von ϱ_{y_0}, und damit ist unsere Behauptung bewiesen.

Jeder Teilraum eines metrischen Raumes ist wieder ein metrischer Raum und daher stets normal. Für beliebige normale Räume ist das allgemein nicht richtig: Ein Teilraum eines normalen Raumes ist nicht notwendig normal. Somit ist die Normalität eines Raumes keine e r b l i c h e Eigenschaft[2]).

Eine erbliche Eigenschaft ist die sogenannte *vollständige Regularität* topologischer Räume, die eine wichtige Verschärfung der Regularität darstellt. Ein T_1-Raum T heißt *vollständig regulär*, wenn für jede abgeschlossene Menge $F \subset T$ und jeden Punkt $x_0 \in T \setminus F$ eine auf T stetige reelle Funktion f mit den Eigenschaften $f(x_0) = 0$, $f(x) = 1$ für $x \in F$ und $0 \leqq f(x) \leqq 1$ existiert. Jeder normale Raum ist vollständig regulär[3]), aber nicht umgekehrt. Jeder Teilraum eines vollständig regulären (insbesondere normalen) Raumes ist also wieder vollständig regulär. A. N. Tichonov, auf den auch der Begriff des vollständig regulären Raumes zurückgeht, zeigte, daß die Klasse der vollständig regulären Räume mit der Klasse aller Teilräume von normalen Räumen übereinstimmt. Vom Standpunkt der Analysis sind vollständig reguläre Räume deshalb wichtig, weil es auf jedem solchen Raum „hinreichend viele" stetige Funktionen gibt, und zwar existiert zu je zwei verschiedenen

[1]) Hier ist $B(x, r)$ wie gewöhnlich die offene Kugel vom Radius r mit dem Mittelpunkt x.

[2]) Eine Eigenschaft eines topologischen Raumes T heißt *erblich*, wenn auch alle Teilräume von T diese Eigenschaft besitzen.

[3]) Diese (durchaus nicht offensichtliche) Tatsache ergibt sich aus dem folgenden Satz von P. S. Uryson: Wenn T ein normaler Raum ist und F_1, F_2 zwei disjunkte abgeschlossene Teilmengen von T sind, dann existiert auf T eine stetige Funktion f, $0 \leqq f(x) \leqq 1$, die auf F_1 gleich 0 und auf F_2 gleich 1 ist.

Punkten x, y eines vollständigen regulären Raumes T eine auf T stetige reelle Funktion, die in diesen Punkten verschiedene Werte annimmt.

2.5.7. Verschiedene Methoden der Angabe einer Topologie in einem Raum. Metrisierbarkeit. In einem Raum können wir eine Topologie erklären, indem wir die Mengen auszeichnen, die wir zu den offenen Mengen rechnen wollen. Die Gesamtheit dieser Mengen muß dabei den Forderungen 1^0 und 2^0 (S. 88) genügen. Gleichwertig damit ist die duale Methode, nämlich die Gesamtheit der abgeschlossenen Mengen anzugeben. Diese Gesamtheit muß dann offenbar den Forderungen 1 und 2 (S. 88) genügen. Tatsächlich kann diese Methode jedoch nur selten angewendet werden. So kann man zum Beispiel schon im Falle der Ebene schwerlich eine unmittelbar Beschreibung aller offenen Teilmengen geben (wie man das für die Zahlengerade machen kann (2.2., Satz 5)).

Eine sehr weit verbreitete Methode der Angabe einer Topologie besteht in der Wahl einer Basis. Die Topologie in metrischen Räumen wird faktisch so eingeführt. Dort gibt man mit Hilfe der Metrik eine Basis an, die Gesamtheit der offenen Kugeln.

Eine weitere mögliche Methode, eine Topologie in einem Raum anzugeben, ist die Einführung eines Konvergenzbegriffes in diesem Raum. Jedoch außerhalb der metrischen Räume ist eine solche Methode nicht allgemein geeignet, da der Übergang einer Menge zu ihrem Abschluß nicht in jedem Fall mit Hilfe konvergenter Folgen beschrieben werden kann, wie sich schon in 2.5.4. zeigte. Diese Methode läßt sich aber universell machen, wenn der Begriff der konvergenten Folge in entsprechender Weise verallgemeinert wird (vgl. etwa [32], Kapitel 2).

Man kann in einem Raum auch dadurch eine Topologie einführen, daß man dort axiomatisch eine Abschließungsoperation definiert. Und zwar nennt man $[\cdot]$ *Abschließungsoperation* in der Menge X, wenn $[\cdot]$ jedem $A \subset X$ eine Menge $[A] \subset X$ zuordnet, so daß die in 2.2.1., Satz 1, angegebenen Eigenschaften 1 bis 4 erfüllt sind. $[A]$ heißt dabei *Abschluß* von A. Erklärt man nun alle Mengen mit $A = [A]$ als abgeschlossen, so ist leicht zu zeigen, daß diese Klasse von Mengen den Bedingungen 1 und 2 (S. 88) genügt, d. h. tatsächlich eine Topologie in X definiert.

Die Angabe einer Metrik ist eine der wichtigsten Methoden der Einführung einer Topologie, wenn sie auch bei weitem nicht universell ist. Wie wir schon sahen, ist jeder metrische Raum normal und genügt dem ersten Abzählbarkeitsaxiom. In Räumen, die auch nur eine dieser zwei Eigenschaften nicht besitzen, kann die Topologie nicht mit Hilfe irgendeiner Metrik definiert werden.

Definition. Ein topologischer Raum heißt *metrisierbar*, wenn man seine Topologie mit Hilfe einer Metrik angeben kann.

Wegen des soeben Gesagten sind die Normalität und die Gültigkeit des ersten Abzählbarkeitsaxioms notwendige Bedingungen für die Metrisierbarkeit eines

Raumes. Andererseits sind diese Bedingungen nicht hinreichend für die Metrisierbarkeit eines Raumes. Jedoch gilt folgender Satz von P. S. Uryson:

Ein topologischer Raum mit abzählbarer Basis ist genau dann metrisierbar, wenn er normal ist.

Die Notwendigkeit dieser Bedingung ist klar, ein Beweis für die Hinlänglichkeit findet sich zum Beispiel in [2].

2.6. Kompaktheit

2.6.1. Der Kompaktheitsbegriff. Eine fundamentale Rolle in der Analysis spielt die folgende Tatsache, die unter dem Namen *Heine-Borelscher Überdeckungssatz* bekannt ist:

Aus jeder Überdeckung des abgeschlossenen Intervalls $[a, b]$ *auf der Zahlengeraden durch offene Intervalle kann man eine endliche Teilüberdeckung auswählen.*

Diese Behauptung bleibt gültig, wenn man anstelle der offenen Intervalle beliebige offene Mengen betrachtet: *Aus jeder offenen Überdeckung des abgeschlossenen Intervalls* $[a, b]$ *kann man eine endliche Teilüberdeckung auswählen.*

Ausgehend von dieser Eigenschaft eines abgeschlossenen Intervalls auf der Zahlengeraden führen wir folgenden wichtigen Begriff ein.

Definition. Ein topologischer Raum T heißt *kompakt*, wenn jede offene Überdeckung von T eine endliche Teilüberdeckung enthält.

Ein kompakter Hausdorffraum heißt *Kompaktum.*

Wie wir unten sehen werden, besitzen neben den abgeschlossenen Intervallen alle abgeschlossenen beschränkten Teilmengen eines euklidischen Raumes von beliebiger endlicher Dimension die Eigenschaft der Kompaktheit. Dagegen sind die Zahlengerade, die Ebene und der dreidimensionale Raum einfachste Beispiele nichtkompakter Räume.

Wir nennen ein System von Teilmengen $\{A\}$ der Menge T *zentriert,* wenn jeder endliche Durchschnitt $\bigcap_{i=1}^{n} A_i$ von Elementen dieses Systems nicht leer ist. Aus der angegebenen Kompaktheitsdefinition und der Dualitätsbeziehung ergibt sich folgender Satz.

Satz 1. *Notwendig und hinreichend für die Kompaktheit eines topologischen Raumes* T *ist folgende Bedingung:*

(R) *Jedes zentrierte System abgeschlossener Teilmengen von* T *besitzt einen nichtleeren Durchschnitt.*

Beweis. $\{F_\alpha\}$ sei ein zentriertes System abgeschlossener Mengen in T, und T sei kompakt. Die Mengen $G_\alpha = T \setminus F_\alpha$ sind dann offen, und kein endliches System

$G_i = T \setminus F_i$ $(i = 1, 2, \ldots, n)$ überdeckt ganz T, da der endliche Durchschnitt $\bigcap_{i=1}^{n} F_i$ nicht leer ist. Dann kann aber auch nicht das System aller G_α die Menge T überdecken (wegen der Kompaktheit, und das bedeutet gerade $\bigcap F_\alpha \neq \emptyset$. Somit ist die Bedingung (R) in T erfüllt, wenn T kompakt ist. Umgekehrt erfülle T die Bedingung (R), und $\{G_\alpha\}$ sei eine offene Überdeckung des Raumes T. Setzen wir $F_\alpha = T \setminus G_\alpha$, so erhalten wir $\bigcap F_\alpha = \emptyset$. Daher kann (wegen der Bedingung (R)) das System $\{F_\alpha\}$ nicht zentriert sein. Also existieren $F_1, \ldots, F_n$ mit $\bigcap_{i=1}^{n} F_i = \emptyset$. Dann bilden aber die entsprechenden $G_i = T \setminus F_i$ eine endliche Teilüberdeckung der Überdeckung $\{G_\alpha\}$. Somit ist die Bedingung (R) gleichwertig mit der Kompaktheit.

Wir führen jetzt einige Grundeigenschaften kompakter Räume an.

Satz 2. *Ist T ein kompakter Raum, so besitzt jede unendliche Teilmenge von T mindestens einen Häufungspunkt.*

Beweis. Wenn T eine unendliche Menge X enthält, die keinen Häufungspunkt besitzt, gibt es in X eine Menge

$$X_1 = (x_1, x_2, \ldots),$$

die ebenfalls keinen Häufungspunkt hat. Dann bilden aber die Mengen

$$X_n = (x_n, x_{n+1}, \ldots), \quad n = 1, 2, \ldots,$$

ein zentriertes System abgeschlossener Mengen in T, die einen leeren Durchschnitt besitzen, d. h., T ist nicht kompakt.

Satz 3. *Eine abgeschlossene Teilmenge eines kompakten Raumes ist kompakt.*

Beweis. F sei eine abgeschlossene Teilmenge des kompakten Raumes T, und $\{F_\alpha\}$ sei ein beliebiges zentriertes System abgeschlossener Teilmengen des Teilraumes $F \subset T$. Dann ist jedes F_α auch abgeschlossen in T, d. h., $\{F_\alpha\}$ ist ein zentriertes System abgeschlossener Mengen in T. Also ist $\bigcap F_\alpha \neq \emptyset$. Nach Satz 1 folgt hieraus die Kompaktheit von F.

Da ein Teilraum eines Hausdorffraumes selbst wieder ein Hausdorffraum ist, erhalten wir hieraus:

Folgerung. *Eine abgeschlossene Teilmenge eines Kompaktums ist ein Kompaktum.*

Satz 4. *Ein Kompaktum K ist in jedem Hausdorffraum T mit $T \supset K$ abgeschlossen.*

Beweis. K sei eine kompakte Menge im Hausdorffraum T, und es sei $y \notin K$. Dann existiert zu jedem Punkt $x \in K$ eine Umgebung U_x des Punktes x und eine Umgebung V_x des Punktes y mit

$$U_x \cap V_x = \emptyset.$$

Die Umgebungen U_x bilden eine offene Überdeckung der Menge K. Wegen der Kompaktheit von K kann man daraus eine endliche Teilüberdeckung $U_{x_1}, U_{x_2}, \ldots, U_{x_n}$ aussondern. Wir setzen nun

$$V = V_{x_1} \cap V_{x_2} \cap \cdots \cap V_{x_n}.$$

Dann ist V eine zu $U_{x_1} \cup U_{x_2} \cup \cdots \cup U_{x_n} \supset K$ disjunkte Umgebung des Punktes y. Folglich ist $y \notin [K]$, und das bedeutet, daß K abgeschlossen ist. Damit ist der Satz bewiesen.

Die Sätze 3 und 4 zeigen, daß die Kompaktheit in der Klasse der Hausdorffräume eine *innere* Eigenschaft des Raumes ist, d. h., jedes Kompaktum bleibt Kompaktum, auch wenn es in einen größeren Hausdorffraum eingebettet wird.

Satz 5. *Jedes Kompaktum ist ein normaler Raum.*

Beweis. Es seien X und Y zwei disjunkte abgeschlossene Teilmengen des Kompaktums K. Durch Wiederholung der Überlegungen aus dem Beweis des vorhergehenden Satzes kann man sich leicht davon überzeugen, daß für einen beliebigen Punkt $y \in Y$ eine Umgebung U_y und eine offene Menge $O_y \supset X$ mit $U_y \cap O_y = \emptyset$ existiert. Damit ist bewiesen, daß jedes Kompaktum regulär ist. Nun durchlaufe y die Menge M. Wir wählen aus der Überdeckung $\{U_y\}$ der Menge Y eine endliche Teilüberdeckung $U_{y_1}, \ldots, U_{y_n}$ aus. Dann genügen die offenen Mengen

$$O^{(1)} = O_{y_1} \cap \cdots \cap O_{y_n} \quad \text{und} \quad O^{(2)} = O_{y_1} \cup \cdots \cup O_{y_n}$$

den Bedingungen

$$O^{(1)} \supset X, \quad O^{(2)} \supset Y \quad \text{und} \quad O^{(1)} \cap O^{(2)} = \emptyset,$$

und das bedeutet gerade Normalität.

2.6.2. Stetige Abbildungen kompakter Räume. Die stetigen Abbildungen eines kompakten Raumes, insbesondere eines Kompaktums, besitzen eine Reihe interessanter und wichtiger Eigenschaften.

Satz 6. *Das stetige Bild eines kompakten Raumes ist ein kompakter Raum.*

Beweis. X sei ein kompakter Raum und f eine stetige Abbildung von X in den topologischen Raum Y. Wir betrachten eine beliebige Überdeckung $\{V_\alpha\}$ des Bildes $f(X)$ durch Mengen, die in $f(X)$ offen sind, und setzen $U_\alpha = f^{-1}(V_\alpha)$. Die Mengen U_α sind offen (als Urbilder offener Mengen bei einer stetigen Abbildung) und bilden eine offene Überdeckung des Raumes X. Aus dieser Überdeckung kann man wegen der Kompaktheit von X eine endliche Teilüberdeckung $U_1, U_2, \ldots, U_n$ auswählen. Dann überdecken aber die Mengen $V_1, V_2, \ldots, V_n$, wobei $V_i = f(U_i)$ ist, das ganze Bild $f(X)$ des Raumes X.

Satz 7. *Eine umkehrbar eindeutige und stetige Abbildung φ eines Kompaktums X in einen Hausdorffraum Y ist ein Homöomorphismus.*

Beweis. Man muß zeigen, daß aus den Bedingungen des Satzes die Stetigkeit der Umkehrabbildung φ^{-1} folgt. Dazu sei F eine abgeschlossene Menge in X und $P \doteq \varphi(F)$ ihr Bild in Y. Nach dem vorhergehenden Satz ist P ein Kompaktum, und folglich ist P abgeschlossen in Y. Also ist das Urbild jeder abgeschlossenen Menge $F \subset X$ bei der Abbildung φ^{-1} abgeschlossen. Das bedeutet aber die Stetigkeit der Abbildung φ^{-1}.

2.6.3. Stetige und halbstetige Funktionen auf kompakten Räumen. Im vorhergehenden Punkt wurden stetige Abbildungen eines Kompaktums in einen Hausdorffraum betrachtet. Einen Spezialfall davon bilden die Abbildungen von einem Kompaktum in die reellen Zahlen, d. h. die reellen Funktionen auf einem Kompaktum. Für diese Funktionen bleiben die grundlegenden Eigenschaften stetiger Funktionen auf einem abgeschlossenen Intervall erhalten, die aus der elementaren Analysis bekannt sind.

Satz 8. *T sei ein kompakter Raum und f eine auf T stetige Funktion. Dann ist f auf T beschränkt und nimmt auf T das Supremum und das Infimum der Funktionswerte an.*

Beweis. Eine stetige Funktion ist eine stetige Abbildung von T in die reellen Zahlen $\boldsymbol{R}^1$. Das Bild von T ist wegen des allgemeinen Satzes 6 kompakt in $\boldsymbol{R}^1$. Wie dem Leser aus der Analysisvorlesung bekannt ist (siehe auch 2.7.2.), ist aber eine kompakte Teilmenge reeller Zahlen beschränkt und abgeschlossen und besitzt daher nicht nur ein endliches Supremum und Infimum, sondern enthält diese Werte auch. Damit ist der Satz bewiesen.

Aufgabe. K sei ein kompakter metrischer Raum und A eine Abbildung von K in sich mit $\varrho(Ax, Ay) < \varrho(x, y)$ für $x \neq y$. Es ist zu zeigen, daß A in K einen einzigen Fixpunkt besitzt.

Die Behauptung des letzten Satzes gestattet eine Verallgemeinerung auch auf eine größere Klasse von Funktionen, und zwar auf die sogenannten halbstetigen Funktionen.

Eine Funktion $f(x)$ heißt *unterhalbstetig* (*oberhalbstetig*) im Punkt x_0, wenn für beliebiges $\varepsilon > 0$ eine Umgebung des Punktes x_0 existiert, in der $f(x) > f(x_0) - \varepsilon$ (bzw. $f(x) < f(x_0) + \varepsilon$) ist.

Zum Beispiel ist die Funktion „ganzer Teil von x“, $f(x) = E(x)$, oberhalbstetig. Wenn wir den Wert $f(x_0)$ einer stetigen Funktion f in irgendeinem Punkt x_0 vergrößern (verkleinern), dann erhalten wir eine oberhalbstetige (unterhalbstetige) Funktion. Ist $f(x)$ oberhalbstetig, so ist $-f(x)$ unterhalbstetig. Diese beiden Bemerkungen gestatten es, sofort eine große Zahl von Beispielen halbstetiger Funktionen zu konstruieren.

Bei der Untersuchung der Eigenschaften der Halbstetigkeit reeller Funktionen ist es bequem, für diese Funktionen auch unendliche Werte zuzulassen. Wenn $f(x_0) = -\infty$ ist, wollen wir die Funktion f im Punkt x_0 stets als unterhalbstetig ansehen; wenn es außerdem für beliebiges $h > 0$ eine Umgebung des Punktes x_0 gibt, in der $f(x) < -h$ ist, nennen wir f auch oberhalbstetig im Punkt x_0.

Wenn $f(x) = +\infty$ ist, dann wollen wir f stets als oberhalbstetig im Punkt x_0 ansehen; wenn es außerdem für beliebiges $h > 0$ eine Umgebung des Punktes x_0 gibt, in der $f(x) > h$ ist, nennen wir f unterhalbstetig im Punkt x_0.

Es sei $f(x)$ eine reelle Funktion auf dem metrischen Raum R. Die (endliche oder unendliche) Größe $\lim\limits_{\varepsilon \to 0} \left[\sup\limits_{x \in B(x_0, \varepsilon)} f(x) \right]$ heißt *oberer Limes* $\bar{f}(x_0)$ der Funktion $f(x)$ im Punkt x_0. Den *unteren Limes*

$\underline{f}(x_0)$ definiert man dann analog, indem man Supremum durch Infimum ersetzt. Die Differenz $\omega f(x_0) = \bar{f}(x_0) - \underline{f}(x_0)$ (sofern sie einen Sinn hat, d. h., wenn $\bar{f}(x_0)$ und $\underline{f}(x_0)$ nicht zugleich $+\infty$ oder zugleich $-\infty$ sind) heißt *Schwankung* der Funktion $f(x)$ im Punkt x_0. Es ist leicht zu sehen, daß eine Funktion $f(x)$ genau dann im Punkt x_0 stetig ist, wenn $\omega f(x_0) = 0$ ist, d. h., wenn $-\infty < \underline{f}(x_0) = \bar{f}(x_0) < \infty$ ist.

Für eine beliebige Funktion $f(x)$ auf einem metrischen Raum ist die Funktion $\bar{f}(x)$ oberhalbstetig und die Funktion $\underline{f}(x)$ unterhalbstetig. Das ergibt sich leicht aus der Definition des oberen und des unteren Limes.

Wir betrachten den metrischen Raum M aller beschränkten reellen Funktionen $\varphi(t)$ auf dem Intervall $[a, b]$. Die Metrik in M definieren wir durch die Gleichung

$$\varrho(x, y) = \varrho(\varphi, \psi) = \sup_{a \leqq t \leqq b} |\varphi(t) - \psi(t)|.$$

Die Funktionen auf M wollen wir *Funktionale* nennen. Das macht man üblicherweise so, um sie von den Funktionen $\varphi(t)$, den Elementen von M, zu unterscheiden.

Wir geben nun ein wichtiges Beispiel eines halbstetigen Funktionals an. Und zwar definieren wir die Länge der Kurve $y = f(x)$ $(a \leqq x \leqq b)$ als das Funktional

$$L_a^b(f) = \sup \sum_{i=1}^{n} \sqrt{(x_i - x_{i-1})^2 + (f(x_i) - f(x_{i-1}))^2},$$

wobei das Supremum (das gleich $+\infty$ sein kann) über alle möglichen Zerlegungen des Intervalls $[a, b]$ genommen wird. Dieses Funktional ist auf dem ganzen Raum M definiert. Für stetige Funktionen stimmt es mit dem Grenzwert

$$\lim_{\max|x_i - x_{i-1}| \to 0} \sum_{i=1}^{n} \sqrt{(x_i - x_{i-1})^2 + (f(x_i) - f(x_{i-1}))^2}$$

überein. Schließlich kann man es für stetig differenzierbare Funktionen in der Form

$$\int_a^b \sqrt{1 + f'^2(x)}\, dx$$

schreiben. Das Funktional $L_a^b(f)$ ist unterhalbstetig in M, was leicht aus seiner Definition folgt. Der obige Satz läßt sich auf halbstetige Funktionen verallgemeinern.

Satz 8a. *Eine unterhalbstetige (oberhalbstetige) endliche Funktion auf einem kompakten T_1-Raum T ist nach unten (oben) beschränkt.*

Um das zu zeigen, nehmen wir an, daß $\inf f(x) = -\infty$ ist. Dann existiert eine Folge $\{x_n\}$ mit $f(x_n) < -n$. Da der Raum T kompakt ist, besitzt die unendliche Teilmenge $\{x_n\}$ (nach Satz 2) wenigstens einen Häufungspunkt x_0. Nach Voraussetzung ist die Funktion f endlich und unterhalbstetig; deshalb findet man eine Umgebung U des Punktes x_0, so daß $f(x) > f(x_0) - 1$ ist für $x \in U$. Damit kann die Umgebung U aber nur endlich viele Punkte der Menge $\{x_n\}$ enthalten, und das widerspricht der Tatsache, daß der Punkt x_0 Häufungspunkt dieser Menge ist. Analog wird der Satz für den Fall einer oberhalbstetigen Funktion bewiesen.

Satz 8b. *Eine unterhalbstetige (oberhalbstetige) endliche Funktion auf einem kompakten T_1-Raum T nimmt ihr Infimum (Supremum) an.*

Die Funktion $f(x)$ sei unterhalbstetig. Dann besitzt sie nach Satz 8a ein endliches Infimum, und es existiert eine Folge $\{x_n\}$ mit $f(x_n) \leqq \inf f(x) + 1/n$.

Weil T kompakt ist, besitzt die Menge $\{x_n\}$ einen Häufungspunkt x_0. Wenn nun $f(x) > \inf f$ wäre, dann könnte man wegen der Unterhalbstetigkeit der Funktion f eine Umgebung U des Punktes x_0 und ein $\delta > 0$ mit $f(x) > \inf f + \delta$ für $x \in U$ finden. Aber dann könnte die Umgebung U keine unendliche Teilmenge der Menge $\{x_n\}$ enthalten. Folglich gilt $f(x_0) = \inf f$, was zu zeigen war.

2.6.4. Abzählbare Kompaktheit

Definition. Ein Raum T heißt *abzählbar-kompakt,* wenn jede unendliche Teilmenge wenigstens einen Häufungspunkt enthält.

Das in 2.6.1., Satz 2, Gezeigte bedeutet, daß jeder kompakte Raum auch abzählbar-kompakt ist. Das Umgekehrte ist allgemein nicht richtig. Dazu geben wir das „traditionelle" Beispiel eines abzählbar-kompakten, aber nicht kompakten Raumes an, und zwar betrachten wir die Menge X aller Ordinalzahlen α, die kleiner als die erste nichtabzählbare Ordinalzahl ω_1 sind. Als *Intervall* (α, β) in X bezeichnen wir die Gesamtheit aller Ordinalzahlen γ, die der Ungleichung $\alpha < \gamma < \beta$ genügen, und als offene Menge in X bezeichnen wir die Vereinigung von beliebig vielen Intervallen. Man kann leicht nachprüfen, daß der konstruierte Raum abzählbar-kompakt, aber nicht kompakt ist.

Die Beziehung zwischen den Begriffen der Kompaktheit und der abzählbaren Kompaktheit wird durch den folgenden Satz geklärt.

Satz 9. *Jede der folgenden zwei Bedingungen ist notwendig und hinreichend dafür, daß der topologische Raum T abzählbar-kompakt ist:*

1. *Jede abzählbare offene Überdeckung des Raumes T enthält eine endliche Teilüberdeckung.*

2. *Jedes abzählbare zentrierte System abgeschlossener Mengen in T besitzt einen nichtleeren Durchschnitt.*

Beweis. Die Gleichwertigkeit der Bedingungen 1 und 2 folgt unmittelbar aus den Dualitätsbeziehungen. Ist nun T nicht abzählbar-kompakt, so ergibt sich genau wie im Beweis von Satz 2, daß in T ein abzählbares zentriertes System abgeschlossener Mengen mit leerem Durchschnitt existiert. Damit ist die Hinlänglichkeit der Bedingung 2 (und damit auch die von 1) festgestellt.

Wir wollen jetzt die Notwendigkeit der Bedingung 2 beweisen. Dazu sei T abzählbar-kompakt und $\{F_n\}$ ein abzählbares zentriertes System abgeschlossener Mengen in T. Dann müssen wir zeigen, daß $\bigcap\limits_n F_n \neq \emptyset$ ist. Setzen wir

$$\Phi_n = \bigcap_{k=1}^{n} F_k,$$

so ist klar, daß alle Φ_n abgeschlossen und (wegen der Zentriertheit von $\{F_n\}$) nicht leer sind. Außerdem gilt offenbar

$$\Phi_1 \supset \Phi_2 \supset \cdots$$

und

$$\bigcap_n \Phi_n = \bigcap_n F_n.$$

Es sind nun zwei Fälle möglich.

1. Von einer gewissen Zahl n_0 an gilt

$$\Phi_{n_0} = \Phi_{n_0+1} = \cdots$$

Dann ist offensichtlich $\bigcap_n \Phi_n = \Phi_{n_0} \neq \emptyset$.

2. Unter den Φ_n gibt es unendlich viele paarweise verschiedene Mengen. Dabei können wir ohne Beschränkung der Allgemeinheit annehmen, daß alle Φ_n voneinander verschieden sind. Es sei

$$x_n \in \Phi_n \setminus \Phi_{n+1}.$$

Die Folge $\{x_n\}$ bildet dann eine unendliche Menge verschiedener Punkte aus T. Wegen der abzählbaren Kompaktheit von T muß sie wenigstens einen Häufungspunkt besitzen, etwa x_0. Weil nun Φ_n alle Punkte $x_n, x_{n+1}, \ldots$ enthält, ist x_0 Häufungspunkt von Φ_n, und wegen der Abgeschlossenheit von Φ_n gilt sogar $x_0 \in \Phi_n$. Folglich ist $x_0 \in \bigcap_n \Phi_n$, d. h. $\bigcap_n \Phi_n \neq \emptyset$.

Auf diese Weise werden sowohl kompakte als auch abzählbar-kompakte Räume durch das „Verhalten" ihrer offenen Mengen charakterisiert. In beiden Fällen kann man aus einer offenen Überdeckung eine endliche Überdeckung auswählen, dabei handelt es sich im ersten Fall um beliebige, im zweiten Fall nur um abzählbare Überdeckungen.

Obwohl im allgemeinen Fall die Kompaktheit nicht aus der abzählbaren Kompaktheit folgt, gilt der folgende Sachverhalt.

Satz 10. *Für einen Raum mit abzählbarer Basis stimmen die Begriffe der Kompaktheit und der abzählbaren Kompaktheit überein.*

Beweis. Wenn nämlich der Raum T eine abzählbare Basis besitzt, dann kann man aus einer beliebigen offenen Überdeckung von T eine abzählbare Teilüberdeckung auswählen (2.5.3., Satz 5). Ist nun T außerdem auch abzählbar-kompakt, so kann man aus dieser abzählbaren Überdeckung wegen des vorhergehenden Satzes eine endliche Teilüberdeckung auswählen. Damit ergibt sich, daß T kompakt ist.

Bemerkung. Der Begriff der abzählbaren Kompaktheit eines topologischen Raumes erscheint (im Gegensatz zur Kompaktheit eines solchen Raumes) nicht sehr natürlich. Er entstand sozusagen „mechanisch". Es zeigt sich aber, daß diese zwei Begriffe für metrische Räume (wie auch für Räume mit abzählbarer Basis) übereinstimmen (wie in 2. 7. bewiesen wird). Für metrische Räume wurde der Begriff der Kompaktheit ursprünglich als die Eigenschaft eingeführt, daß jede unendliche Teilmenge einen Häufungspunkt besitzt, d. h. als abzählbare Kompaktheit. Die „automatische" Übertragung dieser Definition vom metrischen auf den topologischen Fall führte gerade zum Begriff des abzählbar-kompakten topologischen Raumes. Manchmal wird in der Literatur, besonders in der älteren, unter dem Terminus „Kompaktheit" die abzählbare Kompaktheit verstanden. Ein in unserer Terminologie kompakter Raum, in dem man also aus jeder offenen Überdeckung eine endliche

Teilüberdeckung aussondern kann, heißt dann *bikompakt*. Dabei wird ein kompakter Hausdorffraum (d. h. ein Kompaktum) dann *Bikompaktum* genannt, und der Terminus „Kompaktum“ wird für die Bezeichnung eines kompakten *metrischen* Raumes reserviert. Wir werden aber die oben eingeführten Termini (Kompaktheit, abzählbare Kompaktheit) beibehalten. Dabei werden wir natürlich auch kompakte metrische Räume als *Kompaktum* bezeichnen oder, wenn das Vorhandensein einer Metrik ganz besonders betont werden soll, als *metrisches Kompaktum*.

2.6.5. Präkompakte Mengen. Es sei T ein Hausdorffraum und $M \subseteq T$. Ist M nicht abgeschlossen, dann kann M nicht kompakt sein. Zum Beispiel ist eine nichtabgeschlossene Teilmenge der reellen Zahlen nicht kompakt. Es kann sich jedoch herausstellen, daß der Abschluß $[M]$ einer solchen Menge M in T kompakt ist. Dieser Bedingung genügt zum Beispiel jede beschränkte Teilmenge auf der Zahlengeraden oder im n-dimensionalen Raum. Das führt uns zu folgender Definition.

Definition. T sei ein topologischer Raum. Eine Menge M aus T heißt *präkompakt* (oder *kompakt relativ zu* T), wenn ihr Abschluß in T kompakt ist. Analog heißt M *abzählbar-präkompakt* in T, wenn jede unendliche Teilmenge $A \subset M$ wenigstens einen Häufungspunkt besitzt (der nicht notwendig zu M gehören muß).

Der Begriff der Präkompaktheit hängt (im Unterschied zur Kompaktheit) offenbar mit dem Raum T zusammen, in dem wir die gegebene Menge betrachten. So ist zum Beispiel die Menge der rationalen Punkte im Intervall $(0, 1)$ präkompakt, wenn man sie als Teilmenge der Zahlengeraden betrachtet, aber sie ist nicht präkompakt als Teilmenge des Raumes aller rationalen Zahlen.

Der Begriff der Präkompaktheit ist im Fall des metrischen Raumes, um den es im folgenden geht, höchst bedeutsam.

2.7. Kompaktheit in metrischen Räumen

2.7.1. Totalbeschränktheit. Da der metrische Raum einen Spezialfall des topologischen Raumes darstellt, gelten für ihn auch jene Definitionen und Fakten, die in 2.6. dargelegt wurden. Im metrischen Fall hängt die Kompaktheit eng mit dem Begriff der *Totalbeschränktheit* zusammen, den wir jetzt einführen werden.

Es sei M eine Menge im metrischen Raum R und ε eine positive Zahl. Eine Menge A aus R heißt *ε-Netz* für M, wenn es für jeden Punkt $x \in M$ wenigstens einen Punkt $a \in A$ gibt mit

$$\varrho(x, a) \leqq \varepsilon .$$

(Die Menge A braucht nicht notwendig in M enthalten zu sein und braucht sogar keinen einzigen gemeinsamen Punkt mit M zu haben. Jedoch kann man, wenn man

irgendein ε-Netz A für M hat, ein 2ε-Netz $B \subset M$ konstruieren.) Beispielsweise bilden die ganzzahligen Punkte in der Ebene ein $1/\sqrt{2}$-Netz.

Eine Menge M heißt *totalbeschränkt,* wenn für sie zu beliebigem $\varepsilon > 0$ ein endliches ε-Netz existiert. Es ist klar, daß eine totalbeschränkte Menge als Vereinigung endlich vieler beschränkter Mengen notwendigerweise beschränkt ist. Das Umgekehrte ist allgemein nicht richtig, wie das unten angeführte Beispiel 2 zeigt.

Oft ist die folgende triviale Bemerkung nützlich: *Wenn eine Menge M totalbeschränkt ist, dann ist auch ihr Abschluß $[M]$ totalbeschränkt.*

Aus der Definition der Totalbeschränktheit folgt unmittelbar, daß *ein totalbeschränkter metrischer Raum R separabel ist.* Es existiert nämlich für jedes n in R ein endliches $1/n$-Netz, und durch Vereinigung aller dieser Mengen erhält man eine abzählbare Menge M, die überall dicht ist in R. Da ein separabler metrischer Raum eine abzählbare Basis hat (2.5., Satz 4), *besitzt also jeder totalbeschränkte metrische Raum eine abzählbare Basis.*

Beispiele

1. Im n-dimensionalen euklidischen Raum *stimmt die Totalbeschränktheit mit der gewöhnlichen Beschränktheit überein,* d. h. mit der Möglichkeit, eine gegebene Menge in einen hinreichend großen Würfel einzuschließen. Wenn man nämlich einen solchen Würfel in kleine Würfel mit der Kantenlänge ε zerlegt, bilden die Eckpunkte dieser Würfel ein endliches $\sqrt{n}\varepsilon/2$-Netz im ursprünglichen Würfel und folglich erst recht in jeder Menge, die ganz im Innern dieses Würfels liegt.

2. Die Einheitssphäre S im Raum l_2 kann als Beispiel einer beschränkten, aber nicht totalbeschränkten Menge dienen. Dazu betrachten wir in S die Punkte der Form

$$\begin{aligned} e_1 &= (1, 0, 0, \ldots, 0, 0, \ldots), \\ e_2 &= (0, 1, 0, \ldots, 0, 0, \ldots), \\ &\cdots\cdots\cdots\cdots \\ e_n &= (0, 0, 0, \ldots, 1, 0, \ldots), \\ &\cdots\cdots\cdots\cdots \end{aligned}$$

Die Entfernung zwischen zwei beliebigen solchen Punkten e_n und e_m $(n \neq m)$ ist gleich $\sqrt{2}$. Hieraus ist ersichtlich, daß es in S für kein $\varepsilon < \sqrt{2}/2$ ein endliches ε-Netz geben kann.

3. Wir betrachten im l_2 die Menge Π der Punkte

$$x = (x_1, x_2, \ldots, x_n, \ldots),$$

die den Bedingungen

$$|x_1| \leqq 1,\ |x_2| \leqq \frac{1}{2},\ \ldots,\ |x_n| \leqq \frac{1}{2^{n-1}},\ \ldots$$

genügen. Diese Menge heißt *Fundamentalquader* („Hilbertsches Parallelotop") des Raumes l_2. Sie ist Beispiel einer unendlichdimensionalen totalbeschränkten Menge. Zum Beweis ihrer Totalbeschränktheit gehen wir folgendermaßen vor.

Es sei $\varepsilon > 0$ gegeben. Dann wählen wir n so, daß $1/2^{n-1} < \varepsilon/2$ ist. Jedem Punkt

$$x = (x_1, x_2, \ldots, x_n, \ldots) \tag{1}$$

aus Π ordnen wir den Punkt

$$x^* = (x_1, x_2, \ldots, x_n, 0, 0, \ldots) \tag{2}$$

aus derselben Menge zu. Dabei gilt

$$\varrho(x, x^*) = \sqrt{\sum_{k=n+1}^{\infty} x_k^2} \leqq \sqrt{\sum_{k=n}^{\infty} \frac{1}{4^k}} < \frac{1}{2^{n-1}} < \frac{\varepsilon}{2}.$$

Die Menge Π^* aller Punkte der Form (2) aus Π ist totalbeschränkt (als beschränkte Menge im n-dimensionalen Raum). Wählen wir nun in Π^* ein endliches $\varepsilon/2$-Netz aus, so erhalten wir damit offenbar gleichzeitig ein ε-Netz in ganz Π.

2.7.2. Kompaktheit und Totalbeschränktheit

Satz 1. *Ein abzählbar-kompakter metrischer Raum R ist totalbeschränkt.*

Beweis. Wir nehmen an, daß R nicht totalbeschränkt ist, d. h., daß zu jedem $\varepsilon_0 > 0$ in R ein endliches ε_0-Netz existiert. Wir nehmen nun einen beliebigen Punkt a_1 in R. In R gibt es dann wenigstens einen Punkt, den wir mit a_2 bezeichnen, so daß $\varrho(a_1, a_2) > \varepsilon_0$ ist (andernfalls wäre der Punkt a_1 ein ε_0-Netz für R). Weiter gibt es in R einen Punkt a_3 mit

$$\varrho(a_1, a_3) > \varepsilon_0 \quad \text{und} \quad \varrho(a_2, a_3) > \varepsilon_0,$$

sonst wäre das Punktepaar a_1, a_2 ein ε_0-Netz. Wenn allgemein die Punkte $a_1, \ldots, a_k$ bereits konstruiert sind, wählen wir einen Punkt a_{k+1} mit

$$\varrho(a_i, a_{k+1}) > \varepsilon_0, \qquad i = 1, 2, \ldots, k.$$

Diese Konstruktion liefert eine unendliche Folge $a_1, a_2, \ldots$, die wegen $\varrho(a_i, a_j) > \varepsilon_0$ für $i \neq j$ keinen Häufungspunkt besitzt. Dann ist aber T nicht abzählbar-kompakt, was gerade nachzuweisen war.

Somit haben wir gezeigt, daß die abzählbare Kompaktheit in metrischen Räumen die Totalbeschränktheit zur Folge hat, die wiederum die Existenz einer abzählbaren Basis nach sich zieht.

Nach 2.6., Satz 10, erhalten wir folgendes wichtige Resultat.

Folgerung. *Jeder abzählbar-kompakte metrische Raum ist kompakt.*

Wir haben gezeigt, daß die Totalbeschränktheit eine notwendige Bedingung für die Kompaktheit eines metrischen Raumes ist. Diese Bedingung ist aber nicht hinreichend. Zum Beispiel bildet die Gesamtheit der rationalen Punkte des Intervalls $[0, 1]$ mit der gewöhnlichen Abstandsdefinition einen totalbeschränkten, aber nicht kompakten Raum. Denn die Folge

$$0;\ 0{,}4;\ 0{,}41;\ 0{,}414;\ 0{,}4142;\ \ldots$$

von Punkten dieses Raumes, d. h. die Folge der Anfangsabschnitte bei der Dezimalbruchentwicklung von $\sqrt{2} - 1$, besitzt in diesem Raum keinen Grenzwert. Es gilt aber der folgende Satz.

Satz 2. *Ein metrischer Raum R ist genau dann kompakt, wenn folgende beiden Bedingungen gelten:*

1. *R ist totalbeschränkt.*
2. *R ist vollständig.*

Beweis. Die Notwendigkeit der Totalbeschränktheit ist bereits festgestellt worden. Die Notwendigkeit der Vollständigkeit ist trivial, denn wenn $\{x_n\}$ eine Fundamentalfolge in R ist, die keinen Grenzwert besitzt, dann kann diese Folge in R auch keinen Häufungspunkt besitzen.

Wenn andererseits R totalbeschränkt und vollständig ist, so werden wir jetzt zeigen, daß R dann auch kompakt ist. Nach der Folgerung aus Satz 1 brauchen wir nur festzustellen, daß R abzählbar-kompakt ist, d. h., daß jede Folge $\{x_n\}$ von Punkten aus R wenigstens einen Häufungspunkt besitzt.

Dazu wählen wir ein 1-Netz in R und legen um jeden Punkt dieser Menge eine abgeschlossene Kugel vom Radius 1. Weil diese Kugeln ganz R überdecken und ihre Anzahl endlich ist, enthält wenigstens eine von ihnen, wir bezeichnen sie mit B_1, eine unendliche Teilfolge $x_1^{(1)}, \ldots, x_n^{(1)}, \ldots$ der Folge $\{x_n\}$. Weiter wählen wir in B_1 ein 1/2-Netz und legen um jeden Punkt dieses Netzes eine abgeschlossene Kugel vom Radius 1/2. Wenigstens eine dieser Kugeln, die wir mit B_2 bezeichnen, enthält eine unendliche Teilfolge $x_1^{(2)}, \ldots, x_n^{(2)}, \ldots$ der Folge $\{x_n^{(1)}\}$. Weiter finden wir eine abgeschlossene Kugel B_3 vom Radius 1/4, deren Mittelpunkt in B_2 liegt und die eine unendliche Teilfolge $x_1^{(3)}, \ldots, x_n^{(3)}, \ldots$ der Folge $\{x_n^{(2)}\}$ enthält usw. Wir betrachten jetzt neben jeder Kugel B_n die abgeschlossene Kugel A_n mit demselben Mittelpunkt und dem doppelten Radius. Es ist leicht zu sehen, daß die Kugeln A_n ineinandergeschachtelt sind. Wegen der Vollständigkeit des Raumes R ist der Durchschnitt $\bigcap\limits_{n=1}^{\infty} A_n$ nicht leer und besteht aus einem Punkt x_0. Dieser Punkt ist Häufungspunkt der Ausgangsfolge $\{x_n\}$, weil jede Umgebung von x_0 eine gewisse Kugel B_k enthält und somit auch die unendliche Teilfolge $\{x_n^{(k)}\}$ der Folge $\{x_n\}$.

2.7.3. Präkompakte Teilmengen in metrischen Räumen. Den Begriff der Präkompaktheit, der in 2.6.5. für Teilmengen eines beliebigen topologischen Raumes eingeführt wurde, ist insbesondere für Teilmengen eines metrischen Raumes anwendbar. Dabei stimmt hier offenbar der Begriff der abzählbaren Präkompaktheit mit dem Begriff der Präkompaktheit überein. Wir erwähnen den folgenden einfachen, aber wichtigen Sachverhalt.

Satz 3. *Eine Menge M in einem vollständigen metrischen Raum R ist genau dann präkompakt, wenn sie totalbeschränkt ist.*

Der Beweis folgt unmittelbar aus Satz 2 und der trivialen Tatsache, daß eine abgeschlossene Teilmenge eines vollständigen metrischen Raumes selbst wieder vollständig ist.

Die Bedeutung dieses Satzes besteht darin, daß man in der Regel leichter die Totalbeschränktheit einer Menge feststellen kann, als unmittelbar ihre Präkompaktheit zu beweisen. Andererseits ist für Anwendungen in der Analysis gewöhnlich die Präkompaktheit wichtig.

2.7.4. Der Satz von Arzelà. Die Frage nach der Kompaktheit einer Menge in einem metrischen Raum ist eine wichtige Aufgabe. Dabei stößt der Versuch, Satz 2 unmittelbar anzuwenden, auf Schwierigkeiten. Deshalb ist es für Mengen in konkreten Räumen nützlich, spezielle Kompaktheits- bzw. Präkompaktheitskriterien anzugeben, die in der Praxis bequemer sind.

Im n-dimensionalen euklidischen Raum ist die Präkompaktheit einer Menge gleichbedeutend mit ihrer Beschränktheit, wie wir gesehen haben. Jedoch für allgemeinere metrische Räume ist das schon nicht mehr richtig.

Einer der wichtigsten metrischen Räume in der Analysis ist der Raum $C[a, b]$. Für seine Teilmengen liefert der sogenannte *Satz von Arzelà* ein wichtiges und oft verwendetes Kriterium der Präkompaktheit. Um diesen Satz zu formulieren, benötigen wir folgende Begriffe. Eine Familie Φ von Funktionen φ, die auf dem Intervall $[a, b]$ definiert sind, heißt *gleichmäßig beschränkt,* wenn eine Zahl K mit

$$|\varphi(x)| < K$$

für alle $x \in [a, b]$ und alle $\varphi \in \Phi$ existiert.

Eine Familie $\Phi = \{\varphi\}$ heißt *gleichgradig stetig,* wenn für jedes $\varepsilon > 0$ ein $\delta > 0$ existiert, so daß

$$|\varphi(x_1) - \varphi(x_2)| < \varepsilon$$

ist für alle x_1 und x_2 aus $[a, b]$ mit $\varrho(x_1, x_2) < \delta$ und für alle $\varphi \in \Phi$.

Satz 4 (Satz von Arzelà). *Eine Familie Φ stetiger Funktionen auf dem Intervall $[a, b]$ ist genau dann präkompakt in $C[a, b]$, wenn sie gleichmäßig beschränkt und gleichgradig stetig ist.*

B e w e i s. N o t w e n d i g k e i t. Die Familie Φ sei präkompakt in $C[a, b]$. Dann existiert nach dem vorangegangenen Satz für jedes positive ε in der Familie Φ ein endliches $\varepsilon/3$-Netz $\varphi_1, \varphi_2, \ldots, \varphi_k$. Jede der Funktionen φ_i ist als stetige Funktion auf einem abgeschlossenen Intervall beschränkt,

$$|\varphi_i(x)| \leqq K_i .$$

Wir setzen nun $K = \max K_i + \varepsilon/3$. Nach Definition eines $\varepsilon/3$-Netzes gibt es dann für jedes $\varphi \in \Phi$ wenigstens ein φ_i mit

$$\varrho(\varphi, \varphi_i) = \max_x |\varphi(x) - \varphi_i(x)| \leqq \frac{\varepsilon}{3}.$$

Folglich ist

$$|\varphi(x)| \leqq |\varphi_i(x)| + \frac{\varepsilon}{3} \leqq K_i + \frac{\varepsilon}{3} \leqq K.$$

Somit ist Φ gleichmäßig beschränkt.

Weiterhin ist jede der Funktionen φ_i aus dem $\varepsilon/3$-Netz stetig und folglich auf $[a, b]$ auch gleichmäßig stetig. Also existiert zu jedem $\varepsilon/3$ ein δ_i, so daß

$$|\varphi_i(x_1) - \varphi_i(x_2)| < \frac{\varepsilon}{3}$$

gilt, wenn $|x_1 - x_2| < \delta_i$ ist.

Wir setzen nun $\delta = \min \delta_i$ und wählen für eine beliebige Funktion $\varphi \in \Phi$ ein φ_i so, daß $\varrho(\varphi, \varphi_i) < \varepsilon/3$ ist. Dann haben wir für $|x_1 - x_2| < \delta$

$$|\varphi(x_1) - \varphi(x_2)| \leqq |\varphi(x_1) - \varphi_i(x_1)| + |\varphi_i(x_1) - \varphi_i(x_2)| + |\varphi_i(x_2) - \varphi(x_2)|$$
$$< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Damit ist die gleichgradige Stetigkeit von Φ gezeigt.

H i n l ä n g l i c h k e i t. Es sei Φ eine gleichmäßig beschränkte und gleichgradig stetige Familie von Funktionen. Nach Satz 3 können wir die Präkompaktheit von Φ in $C[a, b]$ feststellen, indem wir zeigen, daß zu jedem $\varepsilon > 0$ in $C[a, b]$ ein endliches ε-Netz für Φ existiert. Es sei

$$|\varphi(x)| \leqq K$$

für alle $\varphi \in \Phi$, und es sei $\delta > 0$ so gewählt, daß

$$|\varphi(x_1) - \varphi(x_2)| < \frac{\varepsilon}{5} \quad \text{für} \quad |x_1 - x_2| < \delta$$

und für alle $\varphi \in \Phi$ ist. Wir zerlegen nun das Intervall $[a, b]$ auf der x-Achse durch die Punkte $x_0 = a < x_1 < \cdots < x_n = b$ in Intervalle, deren Länge kleiner als δ ist,

und errichten in diesen Punkten vertikale Geraden. Das Intervall $[-K, K]$ auf der y-Achse zerlegen wir durch die Punkte $y_0 = -K < y_1 < y_2 < \cdots < y_m = K$, deren Abstände weniger als $\varepsilon/5$ betragen, und legen durch die Zerlegungspunkte horizontale Geraden. Auf diese Weise wird das Rechteck $a \leqq x \leqq b$, $-K \leqq y \leqq K$ in „Zellen" zerlegt, deren horizontale Seiten kleiner als δ und deren vertikale Seiten kleiner als $\varepsilon/5$ sind. Jetzt ordnen wir jeder Funktion $\varphi \in \Phi$ einen Polygonzug $\psi(x)$ zu, dessen Ecken Gitterpunkte (x_k, y_l) sind und der sich in den Punkten x_k von der Funktion $\varphi(x)$ um weniger als $\varepsilon/5$ unterscheidet (die Existenz eines solches Polygonzuges ist trivial).

Da nach Konstruktion

$$|\varphi(x_k) - \psi(x_k)| < \frac{\varepsilon}{5},$$

$$|\varphi(x_{k+1}) - \psi(x_{k+1})| < \frac{\varepsilon}{5}$$

und

$$|\varphi(x_k) - \varphi(x_{k+1})| < \frac{\varepsilon}{5}$$

ist, gilt

$$|\psi(x_k) - \psi(x_{k+1})| < \frac{3\varepsilon}{5}.$$

Weil die Funktion $\psi(x)$ zwischen den Punkten x_k und x_{k+1} linear ist, folgt

$$|\psi(x_k) - \psi(x)| < \frac{3\varepsilon}{5} \quad \text{für alle} \quad x \in [x_k, x_{k+1}].$$

Es sei jetzt x ein beliebiger Punkt des Intervalls $[a, b]$ und x_k der nächste links von x gelegene Zerlegungspunkt. Dann ist

$$|\varphi(x) - \psi(x)| \leqq |\varphi(x) - \varphi(x_k)| + |\varphi(x_k) - \psi(x_k)| + |\psi(x_k) - \psi(x)| \leqq \varepsilon.$$

Folglich bilden die Polygonzüge $\psi(x)$ ein ε-Netz für Φ. Ihre Zahl ist offenbar endlich, somit ist Φ totalbeschränkt. Damit ist der Satz vollständig bewiesen.

2.7.5. Der Satz von Peano. Wir zeigen jetzt, wie der Satz von Arzelà zum Beweis des folgenden Existenzsatzes für gewöhnliche Differentialgleichungen mit stetiger rechter Seite angewendet werden kann.

Satz 5 (Peano). *Gegeben sei die Differentialgleichung*

$$\frac{dy}{dx} = f(x, y). \tag{3}$$

Ist die Funktion f in einem abgeschlossenen Gebiet G stetig, so geht durch jeden inneren Punkt (x_0, y_0) dieses Gebietes wenigstens eine Lösungskurve der gegebenen Differentialgleichung.

Beweis. Da die Funktion f in einem abgeschlossenen Gebiet stetig ist, ist sie auch beschränkt,

$$|f(x, y)| < M = \text{const}.$$

Wir legen durch den Punkt (x_0, y_0) die Geraden mit den Anstiegskoeffizienten M und $-M$. Außerdem legen wir zwei vertikale Geraden $x = a$ und $x = b$ so, daß die beiden durch sie abgetrennten Dreiecke mit dem gemeinsamen Eckpunkt (x_0, y_0) ganz innerhalb des Gebietes G liegen. Dieses Paar von Dreiecken bildet eine abgeschlossene Menge, die wir mit Δ bezeichnen.

Wir konstruieren nun sogenannte Eulersche Polygonzüge für die gegebene Gleichung. Dazu legen wir durch den Punkt (x_0, y_0) eine Gerade mit dem Anstiegskoeffizienten $f(x_0, y_0)$. Auf dieser Geraden wählen wir einen Punkt (x_1, y_1) mit $x_1 > x_0$ und legen durch ihn eine Gerade mit dem Anstiegskoeffizienten $f(x_1, y_1)$. Auf dieser Geraden wählen wir einen Punkt (x_2, y_2) mit $x_2 > x_1$, legen durch ihn die Gerade mit dem Anstiegskoeffizienten $f(x_2, y_2)$ usw. Das gleiche Verfahren wird auf eine Folge von Punkten $x_1 < x_0$, $x_2 < x_1$ usw. angewendet. Wir betrachten jetzt eine Folge von Eulerschen Polygonzügen $L_1, L_2, \ldots, L_n, \ldots$, die alle durch den Punkt (x_0, y_0) gehen, wobei die Länge des größten Geradenstückes von L_k für $k \to \infty$ gegen Null strebt. Es sei φ_k die Funktion, deren Bildkurve gerade der Polygonzug L_k ist. Die Funktionen $\varphi_1, \varphi_2, \ldots, \varphi_k, \ldots$ besitzen dann folgende Eigenschaften:

1. Sie sind auf demselben Intervall $[a, b]$ definiert.
2. Sie sind gleichmäßig beschränkt.
3. Sie sind gleichgradig stetig.

Auf Grund des Satzes von Arzelà kann man aus der Folge $\{\varphi_k\}$ eine gleichmäßig konvergente Teilfolge auswählen, etwa die Teilfolge $\varphi^{(1)}, \varphi^{(2)}, \ldots, \varphi^{(k)}, \ldots$

Wir setzen nun

$$\varphi(x) = \lim_{k\to\infty} \varphi^{(k)}(x).$$

Offenbar ist dann $\varphi(x_0) = y_0$. Es bleibt nachzuprüfen, daß φ auf dem Intervall $[a, b]$ die gegebene Differentialgleichung erfüllt. Wir müssen also zeigen, daß für beliebiges $\varepsilon > 0$ die Abschätzung

$$\left|\frac{\varphi(x'') - \varphi(x')}{x'' - x'} - f(x', \varphi(x'))\right| < \varepsilon$$

gilt, sobald $|x'' - x'|$ hinreichend klein ist. Dazu wiederum müssen wir nachweisen, daß für genügend großes k

$$\left|\frac{\varphi^{(k)}(x'') - \varphi^{(k)}(x')}{x'' - x'} - f(x', \varphi^{(k)}(x'))\right| < \varepsilon$$

ist, sobald der Abstand $|x'' - x'|$ hinreichend klein ist.

Da f im Gebiet G stetig ist, gibt es zu jedem $\varepsilon > 0$ ein $\eta > 0$, so daß

$$f(x', y') - \varepsilon < f(x, y) < f(x', y') + \varepsilon \qquad (y' = \varphi(x'))$$

ist für

$$|x - x'| < 2\eta \quad \text{und} \quad |y - y'| < 4M\eta.$$

Die Gesamtheit der Punkte $(x, y) \in G$, die diesen zwei Ungleichungen genügen, bildet ein Rechteck Q. Es sei jetzt K so groß, daß für alle $k > K$

$$|\varphi(x) - \varphi^{(k)}(x)| < 2M\eta$$

ist, und die Länge aller Geradenstücke des Polygonzuges L_k sei kleiner als η. Dann liegen für $|x - x'| < 2\eta$ alle Eulerschen Polygonzüge $\varphi^{(k)}$ mit $k > K$ ganz in Q.

Ferner seien $(a_0, b_0), (a_1, b_1), \ldots, (a_{n+1}, b_{n+1})$ die Eckpunkte des Polygonzuges L_k mit

$$x_0 \leqq a_0 \leqq x' < a_1 < a_2 < \cdots < a_n < x'' \leqq a_{n+1}$$

(wir haben hier $x_0 \leqq x' < x''$ angenommen; die anderen möglichen Fälle werden analog behandelt). Dann folgt für die entsprechenden Funktionen $\varphi^{(k)}$

$$\begin{aligned}
\varphi^{(k)}(a_1) - \varphi^{(k)}(x') &= f(a_0, b_0)\,(a_1 - x'),\\
\varphi^{(k)}(a_{i+1}) - \varphi^{(k)}(a_i) &= f(a_i, b_i)\,(a_{i+1} - a_i), \qquad i = 1, 2, \ldots, n-1,\\
\varphi^{(k)}(x'') - \varphi^{(k)}(a_n) &= f(a_n, b_n)\,(x'' - a_n).
\end{aligned}$$

Hieraus erhalten wir für $|x'' - x'| < \eta$

$$\begin{aligned}
&[f(x', y') - \varepsilon]\,(a_1 - x') < \varphi^{(k)}(a_1) - \varphi^{(k)}(x') < [f(x', y') + \varepsilon]\,(a_1 - x'),\\
&[f(x', y') - \varepsilon]\,(a_{i+1} - a_i) < \varphi^{(k)}(a_{i+1}) - \varphi^{(k)}(a_i)\\
&\qquad\qquad < [f(x', y') + \varepsilon]\,(a_{i+1} - a_i), \qquad i = 1, \ldots, n-1,\\
&[f(x', y') - \varepsilon]\,(x'' - a_n) < \varphi^{(k)}(x'') - \varphi^{(k)}(a_n) < [f(x', y') + \varepsilon]\,(x'' - a_n).
\end{aligned}$$

Summieren wir nun diese Ungleichungen, so ergibt sich

$$[f(x', y') - \varepsilon]\,(x'' - x') < \varphi^{(k)}(x'') - \varphi^{(k)}(x') < [f(x', y') + \varepsilon]\,(x'' - x'),$$

was zu zeigen war.

Verschiedene Teilfolgen von Eulerschen Polygonzügen können gegen verschiedene Lösungen der Gleichung (3) konvergieren. Deswegen ist die Lösungskurve der Gleichung $y' = f(x, y)$, die durch einen Punkt (x_0, y_0) geht, allgemein nicht eindeutig bestimmt.

2.7.6. Gleichmäßige Stetigkeit. Stetige Abbildungen metrischer Kompakta. Für eine Abbildung eines metrischen Raumes in einen metrischen Raum, also insbesondere auch für reelle Funktionen auf metrischen Räumen, ist neben der Stetigkeit auch der für die Analysis wichtige Begriff der gleichmäßigen Stetigkeit von Bedeutung. Eine Abbildung F eines metrischen Raumes X in einen metrischen Raum Y heißt *gleichmäßig stetig*, wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so daß $\varrho_2(F(x_1), F(x_2)) < \varepsilon$ ist für $\varrho_1(x_1, x_2) < \delta$ (ϱ_1 ist hier die Metrik in X und ϱ_2 die Metrik in Y), wobei δ nur von ε, nicht aber von x_1 und x_2 abhängt.

Aufgabe. Es ist zu zeigen, daß die reelle Funktion $F(x) = \sup\limits_{a \leqq t \leqq b} x(t)$ auf dem Raum $C[a, b]$ gleichmäßig stetig ist.

Für stetige Abbildungen metrischer Kompakta gilt der folgende Satz, der den aus der Analysisgrundvorlesung wohlbekannten Satz über stetige Funktionen auf abgeschlossenen Intervallen verallgemeinert.

Satz 6. *Eine stetige Abbildung eines metrischen Kompaktums in einen metrischen Raum ist gleichmäßig stetig.*

Beweis. Die Abbildung F des metrischen Kompaktums K in den metrischen Raum M sei stetig, aber nicht gleichmäßig stetig. Dann gibt es also für ein $\varepsilon > 0$

und jede natürliche Zahl n zwei Punkte x_n und x_n' in K, so daß

$$\varrho_1(x_n, x_n') < \frac{1}{n} \quad \text{und gleichzeitig} \quad \varrho_2\big(F(x_n), F(x_n')\big) \geqq \varepsilon$$

gilt (ϱ_1 ist die Metrik in K, ϱ_2 die Metrik in M). Aus der Folge $\{x_n\}$ kann man wegen der Kompaktheit von K eine Teilfolge $\{x_{n_k}\}$ auswählen, die gegen einen Punkt $x \in K$ konvergiert. Dann konvergiert auch $\{x'_{n_k}\}$ gegen x. Dabei muß für jedes k wenigstens eine der Ungleichungen

$$\varrho_2\big(F(x), F(x_{n_k})\big) \geqq \frac{\varepsilon}{2}, \qquad \varrho_2\big(F(x), F(x'_{n_k})\big) \geqq \frac{\varepsilon}{2}$$

erfüllt sein, was der Stetigkeit der Abbildung F im Punkt x widerspricht.

2.7.7. Der verallgemeinerte Satz von Arzelà. Es seien X und Y zwei metrische Kompakta, und es sei C_{XY} die Menge aller stetigen Abbildungen f des Kompaktums X in Y. In C_{XY} führen wir mit Hilfe der Formel

$$\varrho(f, g) = \sup_{x \in X} \varrho\big(f(x), g(x)\big)$$

einen Abstand ein.

Man prüft leicht nach, daß C_{XY} auf diese Weise zu einem metrischen Raum wird.

Satz 7 (Verallgemeinerter Satz von Arzelà). *Eine Menge $D \subset C_{XY}$ ist genau dann präkompakt, wenn alle Funktionen f auf D gleichgradig stetig sind.*

Das bedeutet, daß zu beliebigem $\varepsilon > 0$ ein $\delta > 0$ existieren soll, so daß aus

$$\varrho(x', x'') < \delta \tag{4}$$

die Beziehung

$$\varrho\big(f(x'), f(x'')\big) < \varepsilon \tag{5}$$

für alle f aus D und x', x'' aus X folgt.

Beweis. Die Notwendigkeit wird ebenso wie in Satz 4 bewiesen.

Wir zeigen nun die Hinlänglichkeit. Dazu betten wir C_{XY} in den Raum M_{XY} aller Abbildungen des Kompaktums X in das Kompaktum Y ein, wobei wir die Metrik

$$\varrho(f, g) = \sup_{x \in X} \varrho\big(f(x), g(x)\big)$$

von C_{XY} auf ganz M_{XY} fortsetzen, und beweisen die Präkompaktheit der Menge D in M_{XY}. Weil C_{XY} in M_{XY} abgeschlossen ist,[1]) folgt aus der Präkompaktheit der Menge D in M_{XY} dann auch ihre Präkompaktheit in C_{XY}.

[1]) Der Grenzwert einer gleichmäßig konvergenten Folge stetiger Abbildungen ist nämlich ebenfalls eine stetige Abbildung. Der erwähnte Satz stellt die unmittelbare Verallgemeinerung eines wohlbekannten Satzes der Analysis dar und wird ebenso wie dieser Satz bewiesen.

Wir geben uns $\varepsilon > 0$ beliebig vor und wählen $\delta > 0$ so, daß aus (4) die Beziehung (5) für alle f aus D und alle x', x'' aus X folgt. Wie man leicht sieht, kann man X als endliche Vereinigung von paarweise disjunkten Mengen E_i darstellen, für deren Punkte jeweils die Beziehung (4) gilt. Dazu braucht man nur die Punkte $x_1, x_2, \ldots, x_n$ eines $\delta/2$-Netzes von X zu nehmen und etwa

$$E_i = B\left(x_i, \frac{\delta}{2}\right) \setminus \bigcup_{j<i} B\left(x_j, \frac{\delta}{2}\right)$$

zu setzen; dabei ist $B\left(x_i, \frac{\delta}{2}\right)$ die Kugel mit dem Mittelpunkt x_i und dem Radius $\frac{\delta}{2}$.

Wir betrachten jetzt im Kompaktum Y irgendein endliches ε-Netz $y_1, y_2, \ldots, y_m$. Mit L bezeichnen wir die Gesamtheit der Funktionen $g(x)$, die auf den Mengen E_i die Werte y_i annehmen. Die Menge dieser Funktionen ist offenbar endlich. Wir wollen zeigen, daß sie ein 2ε-Netz für D in M_{XY} bildet. Dazu betrachten wir eine beliebige Funktion $f \in D$. Für jeden Punkt x_i aus $x_1, \ldots, x_n$ gibt es dann einen Punkt y_j aus $y_1, \ldots, y_m$ mit

$$\varrho(f(x_i), y_j) < \varepsilon.$$

Die Funktion $g \in L$ sei nun so gewählt, daß $g(x_i) = y_j$ ist. Da jeder Punkt $x \in X$ in einer Menge E_i liegt, gilt

$$\begin{aligned}\varrho\big(f(x), g(x)\big) &\leqq \varrho\big(f(x), f(x_i)\big) + \varrho\big(f(x_i), g(x_i)\big) + \varrho\big(g(x_i), g(x)\big) \\ &< 2\varepsilon\end{aligned}$$

für alle $x \in X$. Hieraus folgt, daß die endliche Menge L wirklich ein 2ε-Netz für D ist. Somit ist D präkompakt in M_{XY} und folglich auch in C_{XY}.

2.8. Stetige Kurven in metrischen Räumen[1])

Es sei eine stetige Abbildung

$$P = f(t)$$

des Intervalls $a \leqq t \leqq b$ in den metrischen Raum R gegeben. Wenn t das Intervall von a nach b „durchläuft", dann „durchläuft" der entsprechende Punkt P eine „stetige" Kurve im Raum R. Wir müssen im folgenden strenge Definitionen angeben, um den eben angedeuteten Grundgedanken zu präzisieren.

Zunächst ist einmal die Reihenfolge wesentlich, in der ein Punkt die Kurve durchläuft. Ein und dieselbe Menge, die in Abb. 12 dargestellt ist, die aber in verschiedenen Richtungen durchlaufen wird, wie in den Abb. 13 und 14 angedeutet ist, werden wir als verschiedene Kurven ansehen. Als weiteres Beispiel betrachten wir die reelle Funktion auf dem Intervall [0, 1], die in Abb. 15 dargestellt ist. Sie definiert eine „Kurve", die auf dem Intervall [0, 1] der y-Achse

[1]) Dieser Abschnitt hängt nicht mit den folgenden Ausführungen zusammen. Er kann bei ersten Lesen übersprungen werden.

gelegen ist. Diese Kurve ist vom einfach von 0 bis 1 durchlaufenen Intervall $[0, 1]$ zu unterscheiden, da das Intervall $[A, B]$ dreimal durchlaufen wird (zweimal „nach oben" und einmal „nach unten").

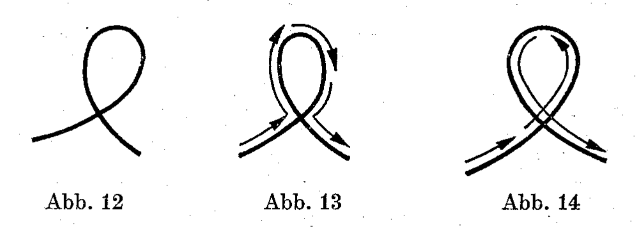


Abb. 12 Abb. 13 Abb. 14

Jedoch werden wir *bei gleichem Durchlaufsinn* der Punkte des Raumes die Auswahl des „Parameters" t als unwesentlich ansehen. Zum Beispiel definieren die in Abb. 15 und 16 dargestellten Funktionen dieselbe auf der y-Achse gelegene Kurve, obwohl die Parameterwerte t, die irgendeinem Punkt der Kurve entsprechen, hier im allgemeinen verschieden sind. So entsprechen etwa dem Punkt A in Abb. 15 auf der t-Achse zwei isolierte Punkte, während ihm in Abb. 16 ein isolierter Punkt und ein rechts davon gelegenes Intervall entsprechen (wenn t dieses Intervall durchläuft, bleibt der Kurvenpunkt an der gleichen Stelle). (Solche Konstanzintervalle des Punktes

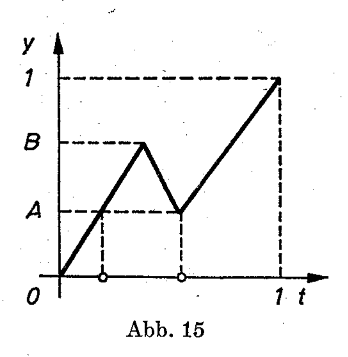


Abb. 15

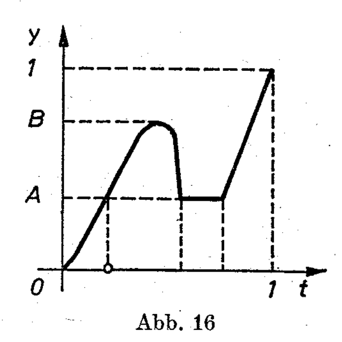


Abb. 16

$P = f(t)$ zuzulassen wird im weiteren bei der Untersuchung der Kompaktheit von Kurvensystemen zweckmäßig sein.)

Wir kommen jetzt zu den formalen Definitionen. Es seien

$$P = f'(t') \quad \text{und} \quad P = f''(t'')$$

zwei stetige Funktionen, die auf den Intervallen

$$a' \leqq t' \leqq b' \quad \text{bzw.} \quad a'' \leqq t'' \leqq b''$$

definiert sind und Werte in einem metrischen Raum R annehmen. Die Funktionen f' und f'' nennen wir *äquivalent*, wenn zwei nichtfallende stetige Funktionen

$$t' = \varphi'(t) \quad \text{und} \quad t'' = \varphi''(t)$$

auf einem Intervall

$$a \leqq t \leqq b$$

mit den Eigenschaften

$$\varphi'(a) = a', \quad \varphi'(b) = b',$$
$$\varphi''(a) = a'', \quad \varphi''(b) = b'',$$
$$f'[\varphi'(t)] = f''[\varphi''(t)]$$

für alle $t \in [a, b]$ existieren.

Es ist leicht zu sehen, daß die so eingeführte Äquivalenzbeziehung reflexiv (f ist äquivalent zu f) und symmetrisch (wenn f' äquivalent ist zu f'', so ist auch f'' zu f' äquivalent) ist. Man kann zeigen, daß sie auch transitiv ist (aus der Äquivalenz von f' und f'' und der Äquivalenz von f'' und f''' folgt die Äquivalenz von f' und f'''). Deshalb zerfallen alle stetigen Funktionen des betrachteten Typs in Klassen zueinander äquivalenter Funktionen. Jede solche Klasse definiert gerade eine *stetige Kurve* im Raum R.

Für eine beliebige Funktion $P = f'(t')$ auf irgendeinem Intervall $[a', b']$ gibt es eine zu ihr äquivalente Funktion auf dem Intervall $[a'', b''] = [0, 1]$. Dazu braucht man nur

$$t' = \varphi'(t) = (b' - a')t + a', \quad t'' = \varphi''(t) = t$$

zu setzen.[1] Somit kann man annehmen, daß jede stetige Kurve in Parameterdarstellung mit Hilfe einer Funktion auf dem Intervall [0, 1] gegeben ist.

Deshalb ist es zweckmäßig, den Raum C_{IR} der stetigen Abbildungen f des Intervalls $I = [0, 1]$ in den Raum R mit der Metrik

$$\varrho(f, g) = \sup_t \varrho(f(t), g(t))$$

zu betrachten.

Wir nennen die Kurvenfolge $L_1, L_2, \ldots, L_n, \ldots$ konvergent gegen die Kurve L, wenn man die Kurven L_n in Parameterdarstellung in der Form

$$P = f_n(t), \quad 0 \leqq t \leqq 1,$$

und die Kurve L in der Form

$$P = f(t), \quad 0 \leqq t \leqq 1,$$

so darstellen kann, daß $\varrho(f, f_n) \to 0$ strebt für $n \to \infty$.

Mit Hilfe des verallgemeinerten Satzes von Arzelà (Satz 7 aus 2.7.) kann man nun leicht den folgenden Satz beweisen.

Satz 1. *Liegt die Folge von Kurven $L_1, L_2, \ldots, L_n, \ldots$ in einem Kompaktum K und sind die Kurven in Parameterdarstellung mit Hilfe von gleichgradig stetigen Funktionen auf dem Intervall* [0, 1] *darstellbar, so kann man aus dieser Folge eine konvergente Teilfolge auswählen.*

Wir definieren jetzt die Länge einer Kurve, die in Parameterdarstellung durch die Funktion

$$P = f(t), \quad a \leqq t \leqq b,$$

gegeben ist, als Supremum aller Summen der Form

$$\sum_{i=1}^{n} \varrho(f(t_{i-1}), f(t_i)),$$

[1]) Wir nehmen dabei an, daß stets $a < b$ ist. Jedoch schließen wir solche „Kurven" nicht aus, die aus einem einzigen Punkt bestehen und die man erhält, wenn die Funktion $f(t)$ auf $[a, b]$ konstant ist. Das wird im weiteren zweckmäßig sein.

wobei die Punkte t_i nur der Bedingung

$$a = t_0 \leqq t_1 \leqq \cdots \leqq t_i \leqq \cdots \leqq t_n = b$$

genügen müssen.

Es ist leicht zu sehen, daß die Länge einer Kurve nicht von der Wahl ihrer Parameterdarstellung abhängt. Wenn man sich bei den Parameterdarstellungen auf Funktionen aus C_{IR} beschränkt, kann man leicht beweisen, daß die Länge einer Kurve ein unterhalbstetiges Funktional bezüglich f (im Raum C_{IR}) ist. In der Sprache der Geometrie kann man dieses Resultat in Form des folgenden Satzes über die Halbstetigkeit ausdrücken.

Satz 2. *Wenn die Kurvenfolge L_n gegen die Kurve L konvergiert, dann ist die Länge von L nicht größer als der untere Limes der Längen der Kurven L_n.*

Wir betrachten jetzt spezielle *Kurven endlicher Länge*. Es sei eine Kurve in Parameterdarstellung durch die Funktion

$$P = f(t), \quad a \leqq t \leqq b,$$

gegeben. Wird die Funktion f nur auf dem Intervall $[a, T]$ mit $a \leqq T \leqq b$ betrachtet, so bestimmt sie einen „Anfangsabschnitt" der Kurve vom Punkt

$$P_a = f(a)$$

zum Punkt

$$P_T = f(T).$$

Dessen Länge sei

$$s = \varphi(T).$$

Man stellt leicht fest, daß

$$P = g(s) = f(\varphi^{-1}(s))$$

eine neue Parameterdarstellung derselben Kurve ist. Jetzt durchläuft s das Intervall $0 \leqq s \leqq S$; dabei ist S die Länge der ganzen betrachteten Kurve. Diese Parameterdarstellung genügt der Bedingung

$$\varrho(g(s_1), g(s_2)) \leqq |s_2 - s_2|$$

(die Länge des Bogens ist nicht kleiner als die Sekante).

Durch Übergang zum Intervall $[0, 1]$ erhalten wir die Parameterdarstellung

$$P = F(\tau) = g(s), \quad \tau = \frac{s}{S},$$

die der Lipschitzbedingung

$$\varrho(F(\tau_1), F(\tau_2)) \leqq S\,|\tau_1 - \tau_2|$$

genügt. *Somit ist für alle Kurven der Länge*

$$S \leqq M,$$

wobei M eine feste Konstante ist, eine gleichgradig stetige Parameterdarstellung mit Hilfe von Funktionen auf dem Intervall $[0, 1]$ *möglich.* Auf diese Funktionen ist folglich Satz 1 anwendbar.

Wir können auf Grund der erhaltenen allgemeinen Resultate zum Beispiel den Beweis des folgenden wichtigen Satzes erbringen.

Satz 3. *Wenn man zwei Punkte A und B eines Kompaktums K innerhalb von K durch eine stetige Kurve endlicher Länge verbinden kann, dann existiert unter diesen Kurven eine kürzeste.*

Mit Y bezeichnen wir das Infimum der Länge aller Kurven, die A und B innerhalb des Kompaktums K verbinden. $L_1, L_2, \ldots, L_n, \ldots$ sei eine Folge von Kurven, die A und B verbinden und deren Längen gegen Y streben. Aus der Folge L_n kann man nach Satz 1 eine konvergente Teilfolge auswählen. Nach Satz 2 kann die Grenzkurve dieser Teilfolge keine Länge besitzen, die größer als Y ist.

Wir bemerken, daß dieser Satz sogar im Fall einer abgeschlossenen glatten (mehrmals differenzierbaren) Hyperebene im dreidimensionalen euklidischen Raum nicht unmittelbar aus den Resultaten einer Vorlesung über Differentialgeometrie folgt, wo man sich gewöhnlich auf den Fall nahe beieinander gelegener Punkte A und B beschränkt.

Alles bisher Angeführte würde größere Durchsichtigkeit erlangen, wenn man die Menge aller Kurven eines gegebenen metrischen Raumes R ebenfalls zu einem metrischen Raum machen könnte. Das ist tatsächlich möglich, wenn man den Abstand zwischen zwei Kurven L_1 und L_2 durch die Formel

$$\varrho(L_1, L_2) = \inf \varrho(f_2, f_2)$$

definiert, wobei das Infimum über alle möglichen Paare von Parameterdarstellungen

$$P = f_1(t), \quad 0 \leqq t \leqq 1,$$

$$\Phi = f_2(t), \quad 0 \leqq t \leqq 1,$$

der Kurven L_1 bzw. L_2 genommen wird.

Der Beweis, daß dieser Abstand den üblichen Axiomen genügt, ist sehr einfach mit Ausnahme eines Punktes. Es bereitet nämlich gewisse Schwierigkeiten, die Identität von zwei Kurven L_1 und L_2 zu beweisen, wenn

$$\varrho(L_1, L_2) = 0$$

ist. Diese Tatsache ist eine unmittelbare Folgerung daraus, daß das Infimum in der Definition des Abstandes $\varrho(L_1, L_2)$ bei entsprechender Wahl der Parameterdarstellungen von f_1 und f_2 angenommen wird. Jedoch ist der Beweis dieser letzten Behauptung auch nicht sehr einfach.

3. Normierte und topologische lineare Räume

3.1. Lineare Räume

Der Begriff des linearen Raumes gehört zu den Grundbegriffen der Mathematik. Er wird nicht nur in diesem Abschnitt, sondern auch in allen weiteren Ausführungen eine wichtige Rolle spielen.

3.1.1. Definition und Beispiele linearer Räume

Definition. Eine nichtleere Menge L von Elementen $x, y, z, \ldots$ heißt *linearer Raum* oder *Vektorraum*, wenn sie folgenden Bedingungen genügt:

I. Für je zwei beliebige Elemente $x, y \in L$ existiert ein eindeutig bestimmtes Element $x + y$ (die *Summe* von x und y), so daß folgende Eigenschaften erfüllt sind:

1. $x + y = y + x$ (Kommutativität).

2. $x + (y + z) = (x + y) + z$ (Assoziativität).

3. In L existiert ein Element 0 mit $x + 0 = x$ für alle $x \in L$ (Existenz des Nullelements).

4. Für jedes $x \in L$ existiert ein Element $-x$ mit $x + (-x) = 0$ (Existenz des inversen Elements).

II. Für jede Zahl α und jedes Element $x \in L$ existiert ein Element $\alpha x \in L$ (das *Produkt* des Elements x mit der Zahl α), so daß folgende Eigenschaften erfüllt sind:

1. $\alpha(\beta x) = (\alpha\beta)x$,

2. $1 \cdot x = x$,

3. $(\alpha + \beta)x = \alpha x + \beta x$,

4. $\alpha(x + y) = \alpha x + \alpha y$.

Je nachdem, welcher Zahlbereich (alle komplexen Zahlen oder nur die reellen Zahlen) verwendet wird, unterscheidet man *komplexe* oder *reelle* lineare Räume.[1]) Wenn nichts Gegenteiliges vorausgesetzt wird, sind unsere Bildungen sowohl für reelle als auch für komplexe Räume richtig.

---

[1]) Man kann auch lineare Räume über beliebigen Körpern betrachten.

Wir bemerken noch, daß jeder komplexe lineare Raum auch als ein reeller Raum angesehen werden kann, wenn man die Skalarmultiplikation von Vektoren auf die reellen Zahlen beschränkt.

Wir betrachten nun Beispiele linearer Räume und überlassen es dem Leser, die oben formulierten Axiome in jedem Fall nachzuprüfen.

1. Die Zahlengerade $\boldsymbol{R}^1$, d. h. die Gesamtheit aller reellen Zahlen mit den gewöhnlichen arithmetischen Operationen der Addition und Multiplikation, bildet einen reellen linearen Raum.

2. Die Gesamtheit aller n-Tupel von reellen Zahlen $x = (x_1, x_2, \ldots, x_n)$ bildet mit der durch die Formeln

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n),$$

$$\alpha(x_1, x_2, \ldots, x_n) = (\alpha x_1, \alpha x_2, \ldots, \alpha x_n)$$

definierten Addition und Multiplikation mit reellen Zahlen ebenfalls einen linearen Raum. Er heißt reeller *n-dimensionaler*[1]) *Raum* und wird mit dem Symbol $\boldsymbol{R}^n$ bezeichnet. Analog wird der komplexe n-dimensionale Raum $\boldsymbol{C}^n$ als die Gesamtheit aller n-Tupel von komplexen Zahlen (mit der Multiplikation mit komplexen Zahlen) definiert.

3. Die stetigen (reellen oder komplexen) Funktionen auf einem Intervall $[a, b]$ bilden mit der üblichen Addition und Skalarmultiplikation von Funktionen den linearen Raum $C[a, b]$, der einen der für die Analysis wichtigsten Räume darstellt.

4. Der Raum l_2 aller Folgen

$$x = (x_1, x_2, \ldots, x_n, \ldots)$$

(reeller oder komplexer) Zahlen, die der Bedingung

$$\sum_{n=1}^{\infty} |x_n|^2 < \infty \tag{1}$$

genügen, bildet mit den Operationen

$$(x_1, x_2, \ldots, x_n, \ldots) + (y_1, y_2, \ldots, y_n, \ldots) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n, \ldots),$$

$$\alpha(x_1, x_2, \ldots, x_n, \ldots) = (\alpha x_1, \alpha x_2, \ldots, \alpha x_n, \ldots)$$

einen linearen Raum. Daß für die Summe zweier Folgen aus l_2 wieder die Bedingung (1) erfüllt ist, folgt aus der elementaren Ungleichung

$$(a_1 + a_2)^2 \leqq 2a_1^2 + 2a_2^2.$$

[1]) Diese Bezeichnungsweise wird im weiteren klar.

5. Die konvergenten Folgen $x = (x_1, x_2, \ldots)$ bilden mit der komponentenweisen Addition und Skalarmultiplikation einen linearen Raum. Wir werden ihn mit c bezeichnen.

6. Die Nullfolgen bilden mit derselben Addition und Skalarmultiplikation ebenfalls einen linearen Raum. Wir werden ihn mit c_0 bezeichnen.

7. Die Gesamtheit m aller beschränkten Zahlenfolgen mit denselben Operationen wie schon in den Beispielen 4 bis 6 stellt einen linearen Raum dar.

8. Schließlich ist auch die Gesamtheit $\boldsymbol{R}^\infty$ aller Zahlenfolgen mit der Addition und der Skalarmultiplikation aus den Beispielen 4 bis 7 ein linearer Raum.

Da die Eigenschaften eines linearen Raumes Eigenschaften der Addition und der Skalarmultiplikation seiner Elemente sind, ist es natürlich, die folgende Definition einzuführen.

Definition 2. Die linearen Räume L und L^* heißen *isomorph*, wenn man zwischen ihren Elementen eine eineindeutige Beziehung herstellen kann, die mit der Operation in L und L^* verträglich ist. Das bedeutet, daß aus

$$x \leftrightarrow x^*$$

und

$$y \leftrightarrow y^*$$

$(x, y \in L;\ x^*, y^* \in L^*)$ die Beziehungen

$$x + y \leftrightarrow x^* + y^*$$

und

$$\alpha x \leftrightarrow \alpha x^*$$

(α beliebige Zahl) folgen.

Isomorphe Räume kann man als verschiedene Realisierungen ein und desselben Raumes ansehen. Als Beispiele isomorpher Räume können der (reelle oder komplexe) n-dimensionale Raum $\boldsymbol{R}^n$ und der Raum aller Polynome (mit reellen bzw. komplexen Koeffizienten), deren Grad $\leqq n - 1$ ist, mit den üblichen Operationen der Addition und Skalarmultiplikation dienen. (Man zeige die Isomorphie!)

3.1.2. Lineare Abhängigkeit. Die Elemente $x, y, \ldots, w$ des lineare Raumes L heißen *linear abhängig*, wenn reelle Zahlen $\alpha, \beta, \ldots, \lambda$ existieren, die nicht alle verschwinden, so daß

$$\alpha x + \beta y + \cdots + \lambda w = 0 \tag{2}$$

ist. Andernfalls heißen diese Elemente *linear unabhängig*. Mit anderen Worten, die Elemente $x, y, \ldots, w$ sind linear unabhängig, wenn aus der Gleichung (2) die Be-

ziehung

$$\alpha = \beta = \cdots = \lambda = 0$$

folgt.

Ein unendliches System $x, y, \ldots$ von Elementen des Raumes L heißt *linear unabhängig*, wenn jedes seiner endlichen Teilsysteme linear unabhängig ist.

Wenn man n linear unabhängige Elemente im Raum L finden kann, aber beliebige $n + 1$ Elemente dieses Raumes linear abhängig sind, sagt man, daß der Raum L die *Dimension* n besitzt. Wenn jedoch in L ein unendliches System von linear unabhängigen Vektoren existiert, sagt man, daß der Raum L *unendlichdimensional* ist. Jedes System von n linear unabhängigen Elementen heißt eine *Basis* im n-dimensionalen Raum L. Die Räume $\boldsymbol{R}^n$ im reellen und $\boldsymbol{C}^n$ im komplexen Fall besitzen die Dimension n, wie man leicht nachprüfen kann, so daß damit ihre Bezeichnung gerechtfertigt ist.

In einer Vorlesung über lineare Algebra werden lineare Räume endlicher Dimension betrachtet. Im Gegensatz dazu werden wir uns im allgemeinen mit unendlichdimensionalen Räumen beschäftigen, die vom Standpunkt der Analysis von grundlegendem Interesse sind. Wir überlassen dem Leser den Nachweis, daß jeder der Räume aus den Beispielen 3 bis 8 unendlichdimensional ist.

3.1.3. Teilräume. Eine nichtleere Teilmenge L' eines linearen Raumes L heißt *Teilraum* oder *Unterraum* von L, wenn sie selbst einen linearen Raum bezüglich der in L definierten Addition und Skalarmultiplikation bildet.

Anders ausgedrückt ist $L' \subset L$ Teilraum von L, wenn aus $x \in L'$, $y \in L'$ auch $\alpha x + \beta y \in L'$ für beliebige α und β folgt.

In jedem linearen Raum L gibt es den Teilraum, der nur aus der Null besteht, den *Nullraum*. Andererseits kann man auch ganz L als Teilraum von L betrachten. Ein Teilraum, der von L verschieden ist und wenigstens ein von Null verschiedenes Element enthält, heißt *echter* Teilraum.

Wir führen jetzt Beispiele echter Teilräume an.

1. Es sei L irgendein linearer Raum mit einem von Null verschiedenen Element x. Dann bildet offenbar die Gesamtheit aller Elemente $\{\lambda x\}$, wobei λ alle (reellen oder komplexen) Zahlen durchläuft, einen eindimensionalen Teilraum. Dies ist ein echter Teilraum, wenn die Dimension von L größer als 1 ist.

2. Wir betrachten den Raum $C[a, b]$ aller stetigen Funktionen (3.1., Beispiel 3) und darin die Gesamtheit aller Polynome $P[a, b]$. Es ist klar, daß die Polynome in $C[a, b]$ einen Teilraum bilden (der, wie auch ganz $C[a, b]$, unendlichdimensional ist). Gleichzeitig kann man den Raum $C[a, b]$ als Teilraum des größeren Raumes aller, stetigen und unstetigen, Funktionen auf $[a, b]$ ansehen.

3. Wir betrachten schließlich die Räume l_2, c_0, c, m und $\boldsymbol{R}^\infty$ (3.1., Beispiele 4 bis 8). Jeder von ihnen bildet einen echten Teilraum des folgenden.

Es sei $\{x_\alpha\}$ eine beliebige nichtleere Menge von Elementen eines linearen Raumes L. Dann existiert in L ein kleinster Teilraum (der möglicherweise mit L übereinstimmt), der $\{x_\alpha\}$ enthält. Es gibt nämlich wenigstens einen Teilraum in L, der $\{x_\alpha\}$ enthält, das ist L selbst. Ferner ist klar, daß *der Durchschnitt einer beliebigen Menge* $\{L_\gamma\}$ *von Teilräumen wieder ein Teilraum ist.* Denn ist $L^* = \bigcap_\gamma L_\gamma$ und $x, y \in L^*$, so folgt $\alpha x + \beta y \in L^*$ für alle α, β. Wir nehmen jetzt alle Teilräume, die das System $\{x_\alpha\}$ von Vektoren enthalten, und betrachten deren Durchschnitt. Das ist gerade der kleinste Teilraum, der das System $\{x_\alpha\}$ enthält. Diesen kleinsten Teilraum nennen wir den *von der Menge* $\{x_\alpha\}$ *erzeugten Teilraum* oder die *lineare Hülle* der Menge $\{x_\alpha\}$ und bezeichnen ihn mit $L(\{x_\alpha\})$.

Aufgaben. Ein linear unabhängiges System $\{x_\alpha\}$ von Elementen eines linearen Raumes L heißt *Hamelbasis*, wenn seine lineare Hülle mit L übereinstimmt. Es sind folgende Behauptungen zu beweisen:

1. In jedem linearen Raum existiert eine Hamelbasis.
Hinweis. Man verwende das Zornsche Lemma.

2. Wenn $\{x_\alpha\}$ eine Hamelbasis in L ist, dann läßt sich jeder Vektor $x \in L$ auf genau eine Weise als endliche Linearkombination von Vektoren aus $\{x_\alpha\}$ darstellen.

3. Je zwei Hamelbasen in einem linearen Raum sind gleichmächtig; die Mächtigkeit einer Hamelbasis eines linearen Raumes nennt man manchmal auch die *algebraische Dimension* dieses Raumes.

4. Lineare Räume sind genau dann isomorph, wenn sie die gleiche algebraische Dimension besitzen.

3.1.4. Faktorräume. Es sei L ein linearer Raum und L' ein Teilraum von L. Zwei Elemente x und y aus L nennen wir *äquivalent*, wenn ihre Differenz $x - y$ zu L' gehört. Diese Relation ist reflexiv, symmetrisch und transitiv, d. h., sie definiert eine Klasseneinteilung aller $x \in L$. Eine Klasse äquivalenter Elemente wird auch *Nebenklasse* (nach dem Teilraum L') genannt. Die Gesamtheit aller dieser Klassen nennen wir *Faktorraum* von L nach L' und bezeichnen ihn mit L/L'.

In jedem Faktorraum kann man auf natürliche Weise eine Addition und eine Skalarmultiplikation einführen. Dazu seinen ξ und η zwei Nebenklassen aus L/L'. Aus jeder dieser Klassen nehmen wir je einen Repräsentanten x bzw. y und nennen die Klasse ζ, die das Element $x + y$ enthält, Summe der Klassen ξ und η. Analog nennen wir die Klasse, die das Element αx enthält, Produkt der Klasse ξ mit der Zahl α. Man kann leicht nachprüfen, daß das Resultat unverändert bleibt, wenn die Repräsentanten x und y durch andere Repräsentanten x' und y' derselben Klassen ξ bzw. η ersetzt werden. Somit haben wir tatsächlich in sinnvoller Weise lineare Operationen für die Elemente des Faktorraumes L/L' definiert. Die unmittelbare Überprüfung zeigt, daß diese Operationen allen Forderungen genügen, die in der Definition eines linearen Raumes vorkommen (man führe den Nachweis durch!). Anders ausgedrückt, *wird jeder Faktorraum* L/L' (mit der soeben definierten Addition und Skalarmultiplikation) *zu einem linearen Raum.*

Wenn L ein n-dimensionaler Raum ist und sein Teilraum L' die Dimension k besitzt, hat der Faktorraum L/L' die Dimension $n - k$ (der Leser möge das nachprüfen!).

Es sei L wieder ein beliebiger linearer Raum und L' ein Teilraum von L. Die Dimension des Faktorraumes L/L' nennt man *Codimension* des Teilraumes L' im Raum L.

Wenn der Teilraum $L' \subset L$ die endliche Codimension n besitzt, kann man in L Elemente $x_1, x_2, \ldots, x_n$ so auswählen, daß jedes Element $x \in L$ (eindeutig) in der Form

$$x = \alpha_1 x_1 + \cdots \alpha_n x_n + y$$

mit Skalaren $\alpha_1, \ldots, \alpha_n$ und $y \in L'$ dargestellt werden kann. Um das zu zeigen, nehmen wir an, der Faktorraum L/L' besitze die Dimension n. Wir wählen dann in diesem Faktorraum eine Basis

$$\xi_1, \xi_2, \ldots, \xi_n$$

und in jeder Klasse ξ_k einen Repräsentanten x_k. Es sei jetzt x ein beliebiges Element aus L und ξ die Klasse in L/L', die x enthält. Dann ist

$$\xi = \alpha_1 \xi_1 + \cdots + \alpha_n \xi_n .$$

Das bedeutet aber nach Definition, daß sich jedes Element aus ξ, also insbesondere x, nur um ein Element aus L' von der entsprechenden Linearkombination der Elemente $x_1, \ldots, x_n$ unterscheidet, d. h., es ist

$$x = \alpha_1 x_1 + \cdots + \alpha_n x_n + y .$$

Den Beweis der Eindeutigkeit dieser Darstellung überlassen wir dem Leser.

3.1.5. Lineare Funktionale. Eine skalare Funktion f auf einem linearen Raum L nennen wir *Funktional*. Ein Funktional f heißt *additiv*, wenn

$$f(x + y) = f(x) + f(y) \qquad \text{für alle } x, y \in L$$

ist, und *homogen*, wenn

$$f(\alpha x) = \alpha f(x)$$

für beliebige Zahlen α ist.

Ein Funktional f auf einem komplexen linearen Raum heißt *antihomogen*, wenn $f(\alpha x) = \bar{\alpha} f(x)$ ist; dabei ist $\bar{\alpha}$ die zu α konjugiert komplexe Zahl.

Ein additives homogenes Funktional heißt *lineares Funktional*. Ein additives antihomogenes Funktional heißt *antilinear*.

Wir geben nun Beispiele linearer Funktionale an.

1. Es sei $a = (a_1, \ldots, a_n)$ ein fixiertes n-Tupel von reellen Zahlen. Durch den Ansatz

$$f(x) = \sum_{i=1}^{n} a_i x_i$$

für jeden Vektor $x = (x_1, \ldots, x_n)$ des reellen n-dimensionalen Raumes $\boldsymbol{R}^n$ wird dann ein lineares Funktional auf dem $\boldsymbol{R}^n$ definiert. Der Ausdruck

$$f(x) = \sum_{i=1}^{n} a_i \bar{x}_i$$

stellt ein antilineares Funktional auf C^n dar. Hierbei können die Zahlen a_i auch komplex sein.

2. Die Integrale

$$I[x] = \int_a^b x(t)\, dt \quad \text{und} \quad \bar{I}[x] = \int_a^b \overline{x(t)}\, dt$$

stellen ein lineares bzw. antilineares Funktional im (reellen bzw. komplexen) Raum $C[a, b]$ dar.

3. Wir betrachten ein allgemeineres Beispiel. Es sei y_0 eine feste stetige Funktion auf $[a, b]$. Dann setzen wir für eine beliebige Funktion $x \in C[a, b]$

$$F(x) = \int_a^b x(t)\, y_0(t)\, dt.$$

Die Linearität dieses Funktionals folgt aus den Grundeigenschaften des Integrals. Das Funktional

$$\bar{F}(x) = \int_a^b \overline{x(t)}\, y_0(t)\, dt$$

ist antilinear (im komplexen Raum $C[a, b]$).

4. Wir betrachten in demselben Raum $C[a, b]$ ein lineares Funktional anderen Typs, und zwar setzen wir

$$\delta_{t_0}(x) = x(t_0).$$

Der Wert des Funktionals δ_{t_0} auf der Funktion $x(t)$ ist also gleich dem Funktionswert an der fixierten Stelle t_0.

Dieses Funktional schreibt man üblicherweise in der Form

$$\delta_{t_0}(x) = \int_a^b x(t)\, \delta(t - t_0)\, dt,$$

wobei man unter δ eine „Funktion" versteht, die überall außer im Punkt 0 verschwindet und deren Integral 1 ergibt (Diracsche δ-Funktion). Solche „Funktionen" erhalten eine strenge Definition im Rahmen der Theorie der Distributionen, deren Grundzüge in 4.4. dargestellt werden.

5. Wir führen noch ein Beispiel eines linearen Funktionals im Raum l_2 an. Es sei k eine fixierte ganze positive Zahl. Dann setzen wir für jedes

$$x = (x_1, \ldots, x_n, \ldots)$$

aus l_2

$$f_k(x) = x_k.$$

Die Linearität eines solchen Funktionals ist trivial. Diese Funktionale gestatten eine „Fortsetzung" auf andere Folgenräume, z. B. auf c_0, c, m, $\boldsymbol{R}^\infty$ (3.1.1, Beispiele 5 bis 8).

3.1.6. Die geometrische Bedeutung eines linearen Funktionals. Es sei f ein nicht identisch verschwindendes lineares Funktional auf einem linearen Raum L. Die Gesamtheit aller Elemente x aus L, die der Bedingung

$$f(x) = 0$$

genügen, bilden einen Teilraum von L. Denn aus $f(x) = f(y) = 0$ folgt auch

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y) = 0.$$

Dieser Teilraum wird *Nullraum* oder *Kern* des Funktionals f genannt und mit Ker f [1]) bezeichnet.

Der Teilraum Ker f besitzt die Codimension 1. Um das zu zeigen, nehmen wir ein Element x_0, das nicht zu Ker f gehört, d. h. ein Element x_0 mit $f(x_0) \neq 0$. Ein solches Element gibt es wegen $f(x) \not\equiv 0$. Ohne Beschränkung der Allgemeinheit kann man annehmen, daß $f(x_0) = 1$ ist, denn andernfalls könnten wir x_0 durch $\dfrac{x_0}{f(x_0)}$ ersetzen. (Es ist klar, daß $f\left(\dfrac{x_0}{f(x_0)}\right) = 1$ ist.) Für jedes Element x setzen wir nun $y = x - f(x)x_0$, dann ist

$$f(y) = f(x - f(x)x_0) = 0,$$

d. h. $y \in \operatorname{Ker} f$. Die Darstellung des Elements x in der Form

$$x = \alpha x_0 + y \quad \text{mit} \quad y \in \operatorname{Ker} f$$

ist bei fixiertem x_0 eindeutig. Wenn nämlich

$$x = \alpha x_0 + y, \qquad y \in \operatorname{Ker} f,$$
$$x = \alpha' x_0 + y', \qquad y' \in \operatorname{Ker} f,$$

[1]) Vom englischen Wort kernel — Kern.

zwei solche Darstellungen sind, so folgt

$$(\alpha - \alpha')\, x_0 = y' - y,$$

und das kann nur für $\alpha = \alpha'$ und damit auch $y = y'$ gelten. Andernfalls würde sich $x_0 = \dfrac{y' - y}{\alpha - \alpha'} \in \operatorname{Ker} f$ ergeben, was der Wahl von x_0 widerspricht.

Hieraus folgt, daß zwei Elemente x_1 und x_2 dann und nur dann zu derselben Nebenklasse nach dem Teilraum $\operatorname{Ker} f$ gehören, wenn $f(x_1) = f(x_2)$ ist. Aus

$$x_1 = f(x_1)\, x_0 + y_1,$$

$$x_2 = f(x_2)\, x_0 + y_2$$

ergibt sich nämlich

$$x_1 - x_2 = \big(f(x_1) - f(x_2)\big)\, x_0 + (y_1 - y_2),$$

und daraus ist ersichtlich, daß $x_1 - x_2 \in \operatorname{Ker} f$ genau dann gilt, wenn $f(x_1) - f(x_2) = 0$ ist.

Jede Nebenklasse ξ nach dem Teilraum $\operatorname{Ker} f$ wird durch einen beliebigen ihrer Repräsentanten bestimmt. Als solchen Repräsentanten kann man nun stets ein Element der Form αx_0 nehmen. Daraus ist ersichtlich, daß der Teilraum $L/\operatorname{Ker} f$ eindimensional ist, d. h., daß $\operatorname{Ker} f$ die Codimension 1 besitzt.

Der Teilraum $\operatorname{Ker} f$ bestimmt umgekehrt ein lineares Funktional, das auf $\operatorname{Ker} f$ verschwindet, bis auf einen konstanten Faktor. Um das zu zeigen, betrachten wir zwei Funktionale f und g mit demselben Kern, $\operatorname{Ker} f = \operatorname{Ker} g$. Wählen wir ein Element x_0 mit $f(x_0) = 1$, so muß $g(x_0) \neq 0$ sein. Denn es ist

$$x = f(x)\, x_0 + y, \qquad y \in \operatorname{Ker} f = \operatorname{Ker} g,$$

und folglich

$$g(x) = f(x)\, g(x_0) + g(y) = f(x)\, g(x_0),$$

und wenn $g(x_0) = 0$ wäre, dann wäre das Funktional g identisch Null. Aus der Gleichung $g(x) = g(x_0)\, f(x)$ folgt aber gerade die Proportionalität von g und f.

Für jeden Teilraum L' der Codimension 1 kann man auch stets ein Funktional f mit $\operatorname{Ker} f = L'$ angeben. Dazu braucht man nur ein beliebiges Element $x_0 \notin L'$ auszuwählen und jedes Element $x \in L$ in der Form $x = \alpha x_0 + y$ mit $y \in L'$ darzustellen. Eine solche Darstellung ist eindeutig. Wenn wir $f(x) = \alpha$ setzen, erhalten wir ein lineares Funktional f mit $\operatorname{Ker} f = L'$ (man prüfe das nach!).

Es sei L' ein Teilraum der Codimension 1 im linearen Raum L, dann wird jede Nebenklasse des Raumes L nach dem Teilraum L' auch als zum Teilraum L' parallele *Hyperebene* bezeichnet (insbesondere ist der Teilraum L' selbst eine Hyperebene, die den Nullpunkt enthält, d. h. die „durch den Koordinatenanfangspunkt

geht“). Mit anderen Worten ist eine zum Teilraum L' parallele Hyperebene M' eine Menge, die aus L' durch Verschiebung um einen Vektor $x_0 \in L$ entsteht:

$$M' = L' + x_0 = \{y: y = x + x_0, x \in L'\}.$$

Offensichtlich ist $M' = L'$ für $x_0 \in L'$ und $M' \neq L'$ für $x_0 \notin L'$. Ist f ein nichttriviales lineares Funktional, dann bildet die Menge $M_f = \{x: f(x) = 1\}$ eine zum Teilraum $\operatorname{Ker} f$ parallele Hyperebene. (Das folgt daraus, daß man jeden Vektor $x \in M_f$ in der Form $x = x_0 + y$ mit $y \in \operatorname{Ker} f$ darstellen kann, wenn x_0 irgendein Element mit $f(x_0) = 1$ ist.) Ist andererseits M' eine zum Teilraum L' parallele Hyperebene (der Codimension 1), die nicht durch den Koordinatenanfangspunkt geht, dann existiert genau ein lineares Funktional f mit $M' = \{x: f(x) = 1\}$. Ist nämlich $M' = L' + x_0$ mit $x_0 \in L$, dann läßt sich jedes Element $x \in L$ eindeutig in der Form $x = \alpha x_0 + y$ mit einem $y \in L'$ darstellen. Wenn wir nun wie oben $f(x) = \alpha$ setzen, erhalten wir das gesuchte lineare Funktional; die Eindeutigkeit ergibt sich daraus, daß jedes Funktional g mit $g(x) = 1$ für $x \in M'$ der Gleichung $g(y) \equiv 0$ für $y \in L'$ genügt, so daß gilt

$$g(\alpha x_0 + y) = \alpha = f(\alpha x_0 + y).$$

Somit besteht eine eineindeutige Beziehung zwischen allen nichttrivialen linearen Funktionalen auf L und allen Hyperebenen in L, die nicht durch den Koordinatenanfangspunkt gehen.

Aufgabe. Es seien $f, f_1, \ldots, f_n$ lineare Funktionale auf dem linearen Raum L mit der Eigenschaft, daß aus $f_1(x) = \cdots = f_n(x) = 0$ auch $f(x) = 0$ folgt. Dann existieren Konstanten $a_1, \ldots, a_n$, so daß $f(x) = \sum_{k=1}^{n} a_k f_k(x)$ für alle $x \in L$ ist.

3.2. Konvexe Mengen und konvexe Funktionale. Der Satz von Hahn-Banach

3.2.1. Konvexe Mengen und konvexe Körper. Vielen wichtigen Teilgebieten der Theorie der linearen Räume liegt der Begriff der *Konvexität* zugrunde. Er geht auf anschauliche geometrische Vorstellungen zurück, aber gleichzeitig gestattet er auch rein analytische Formulierungen.

Es sei L ein reeller linearer Raum und $x, y \in L$. Die Gesamtheit aller Elemente der Form

$$\alpha x + \beta y \quad \text{mit} \quad \alpha, \beta \geqq 0 \quad \text{und} \quad \alpha + \beta = 1$$

nennen wir *abgeschlossene Strecke* in L, die die Punkte x und y verbindet.

Die Strecke ohne die Endpunkte x und y wird *offene Strecke* genannt.

Eine Menge $M \subset L$ heißt *konvex*, wenn sie mit je zwei Punkten x und y auch die diese Punkte verbindende Strecke enthält.

Offener Kern $J(E)$ einer beliebigen Menge $E \subset L$ heißt die Gesamtheit aller der Punkte x aus E, für die zu beliebigem $y \in L$ ein $\varepsilon = \varepsilon(y) > 0$ existiert, so daß $x + ty \in L$ ist für $|t| < \varepsilon$.

Eine konvexe Menge, deren offener Kern nicht leer ist, heißt *konvexer Körper*.

Beispiele

1. Im dreidimensionalen euklidischen Raum stellen Würfel, Kugeln, Tetraeder und Halbräume konvexe Körper dar. Strecken, Ebenen und Dreiecke sind in demselben Raum zwar konvexe Mengen, aber keine konvexen Körper.

2. Wir betrachten im Raum aller stetigen Funktionen auf dem Intervall $[a, b]$ die Menge der Funktionen, die der Bedingung

$$|f(t)| \leqq 1$$

genügen. Diese Menge ist konvex, denn wenn

$$|f(t))| \leqq 1 \quad \text{und} \quad |g(t)| \leqq 1$$

ist, gilt für $\alpha, \beta \geqq 0$ mit $\alpha + \beta = 1$

$$|\alpha f(t) + \beta g(t)| \leqq \alpha + \beta = 1.$$

Aufgabe. Es ist zu prüfen, ob diese Menge ein konvexer Körper ist.

3. Die Einheitskugel in l_2, d. h. die Gesamtheit aller $x = (x_1, \ldots, x_n, \ldots)$ mit $\sum x_n^2 \leqq 1$, ist ein konvexer Körper. Sein offener Kern besteht aus den Punkten, die der Bedingung $\sum x_n^2 < 1$ genügen.

4. Der Fundamentalquader Π in l_2 ist eine konvexe Menge, aber kein konvexer Körper. Um das zu zeigen, sei $x \in \Pi$; dann gilt also $|x_n| \leqq 1/2^{n-1}$ für alle $n = 1, 2, \ldots$. Außerdem wählen wir $y_0 = (1, 1/2, \ldots, 1/n, \ldots)$. Ist nun $x + ty_0 \in \Pi$, so gilt $|x_n + t/n| \leqq 1/2^{n-1}$. Daraus ergibt sich

$$\left|\frac{t}{n}\right| \leqq \left|x_n + \frac{t}{n}\right| + |x_n| \leqq \frac{1}{2^{n-1}} + \frac{1}{2^{n-1}} = \frac{1}{2^{n-2}},$$

woraus $t = 0$ folgt. Also ist der offene Kern der Menge Π leer.

Aufgabe. Φ sei die Gesamtheit aller Punkte $x = (x_1, \ldots, x_{0_n}, \ldots)$ aus l_2, die der Bedingung $\sum n^2 x_n^2 \leqq 1$ genügen. Man zeige, daß Φ eine konvexe Menge, aber kein konvexer Körper ist.

Ist M eine konvexe Menge, so ist ihr offener Kern $J(M)$ ebenfalls konvex. Dazu betrachten wir $x, y \in J(M)$ und $z = \alpha x + \beta y$, $\alpha, \beta \geqq 0$, $\alpha + \beta = 1$. Dann gibt es zu gegebenem $a \in L$ stets Zahlen $\varepsilon_1 > 0$ und $\varepsilon_2 > 0$, so daß die Punkte $x + t_1 a$ und $y + t_2 a$ für $|t_1| < \varepsilon_1$ bzw. $|t_2| < \varepsilon_2$ zur Menge M gehören. Folglich gehört auch der Punkt $\alpha(x + ta) + \beta(y + ta) = z + ta$ für $|t| < \min(\varepsilon_1, \varepsilon_2)$ zu M, d. h. z zu $J(M)$.

Die konvexen Mengen besitzen folgende wichtige Eigenschaft.

Satz 1. *Der Durchschnitt beliebig vieler konvexer Mengen ist eine konvexe Menge.*

Beweis. Es sei $M = \bigcap_\alpha M_\alpha$, und alle M_α seien konvexe Mengen. Weiter seien x und y zwei beliebige Punkte aus M. Dann gehört die Verbindungsstrecke von x und y zu jedem M_α und folglich auch zu M. Somit ist M konvex.

Wir bemerken, daß der Durchschnitt konvexer Körper (der ja eine konvexe Menge ist) kein konvexer Körper zu sein braucht (man gebe ein Beispiel dafür an).

Für jede Menge A eines linearen Raumes L existiert eine kleinste konvexe Menge, die A enthält, und zwar der Durchschnitt aller konvexen Mengen, die A enthalten (es existiert wenigstens eine konvexe Menge, die A enthält, nämlich ganz L). Diese minimale konvexe Menge, die A enthält, nennen wir *konvexe Hülle* von A.

Zur konvexen Hülle betrachten wir ein wichtiges Beispiel. Es seien $x_1, x_2, \ldots, x_{n+1}$ Punkte eines linearen Raumes. Wir sagen, daß sich diese Punkte *in allgemeiner Lage* befinden, wenn die Vektoren $x_2 - x_1, x_3 - x_1, \ldots, x_{n+1} - x_1$ linear unabhängig sind. (Das ist gleichwertig damit, daß $\sum_{i=1}^{n+1} \lambda_i x_i = 0$ und $\sum_{i=1}^{n+1} \lambda_i = 0$ nur für $\lambda_1 = \lambda_2 = \cdots = \lambda_{n+1} = 0$ gelten kann.) Die konvexe Hülle der Punkte $x_1, x_2, \ldots, x_{n+1}$, die sich in allgemeiner Lage befinden, heißt *n-dimensionales Simplex*. Die Punkte $x_1, x_2, \ldots, x_{n+1}$ sind dabei dessen Eckpunkte. Ein nulldimensionales Simplex ist ein Punkt, ein eindimensionales Simplex eine Strecke, ein zweidimensionales Simplex ein Dreieck und ein dreidimensionales Simplex ein Tetraeder.

Befinden sich die Punkte $x_1, x_2, \ldots, x_{n+1}$ in allgemeiner Lage, so befinden sich jeweils $k + 1$ davon $(k < n)$ ebenfalls in allgemeiner Lage und erzeugen folglich ein k-dimensionales Simplex, eine sogenannte *k-dimensionale Seite* des gegebenen n-dimensionalen Simplexes. Zum Beispiel besitzt das Tetraeder mit den Eckpunkten, e_1, e_2, e_3, e_4 vier zweidimensionale Seiten, die durch die Eckpunkttripel (e_2, e_3, e_4), (e_1, e_3, e_4), (e_1, e_2, e_4) bzw. (e_1, e_2, e_3) bestimmt sind, sechs eindimensionale Seiten und vier nulldimensionale.

Satz 2. *Ein Simplex mit den Eckpunkten $x_1, x_2, \ldots, x_{n+1}$ ist die Gesamtheit aller Punkte, die man in der Form*

$$x = \sum_{k=1}^{n+1} \alpha_k x_k, \quad \alpha_k \geqq 0, \quad \sum_{k=1}^{n+1} \alpha_k = 1 \tag{1}$$

darstellen kann.

Beweis. Man prüft leicht nach, daß die Gesamtheit S aller Punkte der Form (1) eine konvexe Menge darstellt, die die Punkte $x_1, x_2, \ldots, x_{n+1}$ enthält. Andererseits muß jede konvexe Menge, die diese Punkte enthält, auch alle Punkte der Form (1) enthalten. Folglich ist S die kleinste konvexe Menge, die die Punkte $x_1, x_2, \ldots, x_{n+1}$ enthält.

3.2.2. Konvexe Funktionale. Eng verbunden mit dem Begriff der konvexen Menge ist der wichtige Begriff des konvexen Funktionals.

Definition. Ein nichtnegatives Funktional p auf einem reellen linearen Raum L heißt *konvex*, wenn

1. $p(x + y) \leqq p(x) + p(y)$ für alle $x, y \in L$;

2. $p(\alpha x) = \alpha p(x)$ für alle $\alpha > 0$ ist.

Es wird nicht verlangt, daß die Größe $p(x)$ für alle $x \in L$ endlich ist, d. h., der Fall $p(x) = \infty$ ist zugelassen.

Wir geben einige Beispiele konvexer Funktionale an.

1. Der Betrag eines Vektors im n-dimensionalen euklidischen Raum $\boldsymbol{R}^n$. Hier bedeutet die erste Bedingung, daß der Betrag der Summe zweier Vektoren die Summe ihrer Beträge nicht übersteigt (Dreiecksungleichung). Die zweite Bedingung folgt unmittelbar aus der Definition des Betrages in $\boldsymbol{R}^n$.

2. M sei der Raum aller beschränkten Funktionen x auf einer Menge S und s_0 ein fester Punkt in S. Dann ist

$$p_{s_0}(x) = |x(s_0)|$$

ein konvexes Funktional.

3. Es sei m der Raum aller beschränkten Zahlenfolgen $x = (x_1, x_2, \ldots, x_n, \ldots)$. Hier ist das Funktional

$$p(x) = \sup_n |x_n|$$

konvex.

3.2.3. Das Minkowskische Funktional. Wir betrachten jetzt den Zusammenhang zwischen konvexen Funktionalen und konvexen Mengen.

Satz 3. *Ist p ein konvexes Funktional auf einem linearen Raum L und k eine positive Zahl, dann ist die Menge*

$$E = \{x : p(x) \leqq k\}$$

konvex. Ist das Funktional überall endlich, dann bildet E einen konvexen Körper, dessen offener Kern die Menge

$$\{x : p(x) < k\}$$

ist (die offenbar den Punkt 0 enthält).

Beweis. Ist $x, y \in E$ und $\alpha + \beta = 1$, $\alpha, \beta \geqq 1$, so gilt

$$p(\alpha x + \beta y) \leqq \alpha p(x) + \beta p(y) \leqq k,$$

d. h., E ist konvex. Das Funktional p sei jetzt endlich, $p(x) < k$, $t > 0$ und $y \in L$. Dann gilt

$$p(x \pm ty) \leqq p(x) + tp(\pm y).$$

Wenn hier $p(-y) = p(y) = 0$ ist, folgt $x \pm ty \in E$ für alle t; wenn jedoch wenigstens eine der beiden nichtnegativen Zahlen $p(y)$, $p(-y)$ verschieden von Null ist, folgt $x \pm ty \in E$ für

$$t < \frac{k - p(x)}{\max\big(p(y), p(-y)\big)}.$$

Damit ist der Satz bewiesen.

Wir wählen nun einen bestimmten Wert k, etwa $k = 1$. Dann bestimmt jedes endliche konvexe Funktional in L eindeutig den konvexen Körper $E = \{x: p(x) \leqq 1\}$, dessen offener Kern den Punkt 0 enthält. Umgekehrt sei E ein konvexer Körper mit $0 \in J(E)$. Dann ist

$$p_E(x) = \inf\left\{r: \frac{x}{r} \in E, r > 0\right\} \tag{2}$$

ein endliches konvexes Funktional. Es wird *Minkowskisches Funktional* des konvexen Körpers E genannt.

Wir wollen die Konvexität des Minkowskischen Funktionals (2) nachweisen. Für jedes $x \in L$ gehört das Element x/r für hinreichend großes r zu E. Deshalb ist die Größe $p_E(x)$, die durch Gleichung (2) definiert wird, nichtnegativ und endlich. Ist $t > 0$ und $y = tx$, so gilt

$$\begin{aligned} p_E(y) &= \inf\left\{r > 0: \frac{y}{r} \in E\right\} = \inf\left\{r > 0: \frac{tx}{r} \in E\right\} \\ &= \inf\left\{tr' > 0: \frac{x}{r'} \in E\right\} \\ &= t\inf\left\{r' > 0: \frac{x}{r'} \in E\right\} = tp_E(x). \end{aligned} \tag{3}$$

Jetzt seien $x_1, x_2 \in L$ und $\varepsilon > 0$ beliebig. Wenn wir die Zahlen r_i $(i = 1, 2)$ so wählen, daß $p_E(x_i) < r_i < p_E(x_i) + \varepsilon$ ist, dann gehört x_i/r_i zu E. Setzen wir $r = r_1 + r_2$, so liegt $\dfrac{x_1 + x_2}{r} = \dfrac{r_1 x_1}{r r_1} + \dfrac{r_2 x_2}{r r_2}$ auf der Strecke mit den Endpunkten x_1/r_1 und

x_2/r_2. Wegen der Konvexität von E gehört aber diese Strecke und somit auch $\frac{x_1 + x_2}{r}$ zu E, woraus

$$p_E(x_1 + x_2) \leqq r = r_1 + r_2 < p_E(x_1) + p_E(x_2) + 2\varepsilon$$

folgt. Weil hier $\varepsilon > 0$ beliebig war, gilt also

$$p_E(x_1 + x_2) \leqq p_E(x_1) + p_E(x_2). \tag{4}$$

Die Beziehungen (3) und (4) bedeuten aber gerade die Konvexität des Funktionals $p_E(x)$.

Aufgabe. Eine Menge A in einem linearen Raum L heißt *absorbierend*, wenn für jedes $x \in L$ ein $\alpha > 0$ existiert, so daß $x \in \lambda A$ ist für alle $\lambda \geqq \alpha$. Man beweise, daß eine konvexe Menge A genau dann absorbierend ist, wenn ihr Minkowskisches Funktional endlich ist.

3.2.4. Der Satz von Hahn-Banach. Es sei L ein reeller linearer Raum und L_0 ein Teilraum von L. Weiter sei auf dem Teilraum L_0 ein lineares Funktional f_0 gegeben. Ein lineares Funktional f, das auf dem ganzen Raum L definiert ist, heißt *Fortsetzung* des Funktionals f_0, wenn

$$f(x) = f_0(x) \quad \text{für alle } x \in L_0$$

ist. Dem Problem der Fortsetzung eines linearen Funktionals begegnet man oft in der Analysis. Eine fundamentale Rolle in diesem Problemkreis spielt der folgende Satz.

Satz 4 (HAHN-BANACH). *Es sei p ein endliches konvexes Funktional auf einem reellen linearen Raum L und L_0 ein linearer Teilraum von L. Wenn f_0 ein lineares Funktional auf L_0 ist, das auf L_0 durch $p(x)$ majorisiert wird, d. h., wenn auf L_0*

$$f_0(x) \leqq p(x) \tag{5}$$

ist, dann kann f_0 zu einem linearen Funktional f auf L fortgesetzt werden, das durch $p(x)$ auf ganz L majorisiert wird.

Beweis. Wir werden zeigen, daß man das Funktional f_0 im Fall $L_0 \neq L$ von L_0 auf einen größeren Teilraum L' unter Beibehaltung der Bedingung (5) fortsetzen kann. Es sei z ein beliebiges Element aus L, das nicht zu L_0 gehört, und L' der von L_0 und z erzeugte Teilraum. Jedes Element aus L' besitzt also die Form

$$tz + x \quad \text{mit} \quad x \in L_0.$$

Ist f' eine Fortsetzung von f_0 auf L', dann gilt

$$f'(tz + x) = tf'(z) + f_0(x)$$

oder, wenn $f'(z) = c$ gesetzt wird,

$$f'(tz + x) = tc + f_0(x).$$

Wir wollen c so wählen, daß die Bedingung (5) auch auf L' gilt, daß also für alle $x \in L_0$ und alle reellen Zahlen t die Ungleichung

$$f_0(x) + tc \leqq p(x + tz)$$

erfüllt ist. Für $t > 0$ ist das gleichwertig mit der Bedingung

$$f_0\left(\frac{x}{t}\right) + c \leqq p\left(\frac{x}{t} + z\right) \quad \text{bzw.} \quad c \leqq p\left(\frac{x}{t} + z\right) - f_0\left(\frac{x}{t}\right) \tag{6}$$

und für $t < 0$ mit der Bedingung

$$f_0\left(\frac{x}{t}\right) + c \geqq -p\left(-\frac{x}{t} - z\right) \quad \text{bzw.} \quad c \geqq -p\left(-\frac{x}{t} - z\right) - f_0\left(\frac{x}{t}\right). \tag{7}$$

Wir werden zeigen, daß immer eine Zahl c existiert, die diese beiden Bedingungen erfüllt. Dazu seien y' und y'' beliebige Elemente aus L_0. Dann ist

$$-f_0(y'') + p(y'' + z) \geqq -f_0(y') - p(-y' - z). \tag{8}$$

Das folgt aus der Ungleichung

$$\begin{aligned} f_0(y'') - f_0(y') \leqq p(y'' - y') &= p\big((y'' + z) - (y' - z)\big) \\ &\leqq p(y'' + z) + p(-y' - z). \end{aligned}$$

Nun setzen wir

$$c'' = \inf_{y''} \big(-f_0(y'') + p(y'' + z)\big),$$

$$c' = \sup_{y'} \big(-f_0(y') - p(-y' - z)\big).$$

Aus (8) folgt dann $c'' \geqq c'$, da y' und y'' beliebig waren. Wir wählen jetzt c so, daß

$$c'' \geqq c \geqq c'$$

ist, und definieren das Funktional f' auf L' durch die Formel

$$f'(tz + x) = tc + f_0(x).$$

Dieses Funktional genügt dann der Ungleichung (5).

Somit haben wir gezeigt, daß man ein Funktional f_0, das auf einem Teilraum $L_0 \subset L$ definiert ist und auf L_0 der Bedingung (5) genügt, unter Beibehaltung dieser Bedingung auf einen größeren Teilraum L' fortsetzen kann.

Wenn in L ein abzählbares Erzeugendensystem $x_1, x_2, \ldots, x_n, \ldots$ existiert, konstruieren wir das Funktional auf L induktiv, indem wir die aufsteigende Kette von Teilräumen

$$L^{(1)} = \{L_0, x_1\}, L^{(2)} = \{L^{(1)}, x_2\}, \ldots$$

betrachten (hier bedeutet $\{L^{(k)}, x_{k+1}\}$ den minimalen linearen Teilraum in L, der $L^{(k)}$ und x_{k+1} enthält). Dann liegt jedes Element $x \in L$ in einem gewissen $L^{(k)}$, und folglich wird das Funktional auf ganz L fortgesetzt.

Im allgemeinen Fall (d. h., wenn kein endliches Erzeugendensystem von L existiert) wird der Beweis durch Anwendung des Zornschen Lemmas vollendet. Die Gesamtheit $\mathscr{F}$ aller möglichen Fortsetzungen des Funktionals f_0, die der Bedingung (5) genügen, ist halbgeordnet. Dabei besitzt jede linear geordnete Teilmenge $\mathscr{F}_0$ ein Supremum, nämlich das Funktional, das auf der Vereinigung der Definitionsgebiete aller Funktionale $f' \in \mathscr{F}_0$ definiert ist und mit jedem solchen f' auf dessen Definitionsgebiet übereinstimmt. Nach dem Zornschen Lemma existiert dann ein maximales Element f in $\mathscr{F}$. Dieses maximale Element ist ein Funktional mit den gewünschten Eigenschaften. Denn es ist einmal eine Fortsetzung des Ausgangsfunktionals f_0, und zum anderen genügt es der Bedingung (5) auf seinem ganzen Definitionsgebiet. Dieses Definitionsgebiet aber ist ganz L, weil wir f sonst in der oben beschriebenen Weise wiederum auf einen größeren Teilraum fortsetzen könnten und f dann nicht maximal wäre. Damit ist der Satz bewiesen.

Wir geben noch eine komplexe Variante des Satzes von HAHN-BANACH an.

Ein nichtnegatives Funktional p auf einem komplexen linearen Raum L heißt *konvex*, wenn für alle $x, y \in L$ und alle komplexen Zahlen λ gilt

$$p(x + y) \leqq p(x) + p(y),$$

$$p(\lambda x) = |\lambda| \, p(x).$$

Satz 4a. *Es sei p ein endliches konvexes Funktional auf einem komplexen linearen Raum L und f_0 ein lineares Funktional, das auf einem linearen Teilraum $L_0 \subset L$ definiert ist und dort der Bedingung*

$$|f_0(x)| \leqq p(x), \quad x \in L_0,$$

genügt. Dann existiert ein lineares Funktional f auf ganz L mit

$$|f(x)| \leqq p(x), \qquad x \in L,$$

$$f(x) = f_0(x), \qquad x \in L_0.$$

Beweis. Wir bezeichnen mit L_R und L_{0R} die Räume L und L_0, die jetzt als reelle lineare Räume angesehen werden. Offenbar ist dann p ein endliches konvexes Funktional auf L_R und $f_{0R}(x) = \operatorname{Re} f_0(x)$ ein reelles lineares Funktional auf L_{0R}, das der

Bedingung

$$|f_{0R}(x)| \leqq p(x)$$

und damit erst recht der Bedingung

$$f_{0R}(x) \leqq p(x)$$

genügt. Nach Satz 4 existiert also ein reelles lineares Funktional f_R, das auf ganz L_R definiert ist und den Bedingungen

$$f_R(x) \leqq p(x), \qquad x \in L_R(= L),$$

$$f_R(x) = f_{0R}(x), \qquad x \in L_{0R}(= L_0),$$

genügt. Offenbar ist $-f_R(x) = f_R(-x) \leqq p(-x) = p(x)$, so daß also

$$|f_R(x)| \leqq p(x), \qquad x \in L_R(= L), \tag{9}$$

gilt. Wir definieren nun f auf L durch den Ansatz

$$f(x) = f_R(x) - i f_R(ix)$$

(hierbei benutzen wir, daß L ein komplexer linearer Raum ist, so daß in L die Multiplikation mit einer komplexen Zahl erklärt ist). Die unmittelbare Überprüfung zeigt, daß f ein komplexes lineares Funktional auf L ist mit

$$f(x) = f_0(x) \quad \text{für} \quad x \in L_0,$$

$$\operatorname{Re} f(x) = f_R(x) \quad \text{für} \quad x \in L.$$

Zu zeigen bleibt noch, daß $|f(x)| \leqq p(x)$ ist für alle $x \in L$. Dazu nehmen wir das Gegenteil an. Dann gilt $|f(x_0)| > p(x_0)$ für ein gewisses $x_0 \in L$. Die komplexe Zahl $f(x_0)$ stellen wir in der Form $f(x_0) = \varrho e^{i\varphi}$ mit $\varrho > 0$ dar und setzen $y_0 = e^{-i\varphi} x_0$. Dann ist

$$f_R(y_0) = \operatorname{Re} f(y_0) = \operatorname{Re} [e^{-i\varphi} f(x_0)] = \varrho > p(x_0) = p(y_0),$$

was der Bedingung (9) widerspricht. Damit ist der Satz bewiesen.

Aufgabe. Man zeige, daß die Forderung der Endlichkeit des Funktionals p im Satz von HAHN-BANACH weggelassen werden kann.

3.2.5. Die Trennbarkeit konvexer Mengen in einem linearen Raum. Es seien L ein reeller Raum und M und N zwei Teilmengen von L. Man sagt, daß ein auf L definiertes lineares Funktional f diese Mengen *trennt*, wenn eine Zahl C existiert mit

$$f(x) \geqq C \quad \text{für} \quad x \in M \quad \text{und} \quad f(x) \leqq C \quad \text{für} \quad x \in N.$$

Die beiden folgenden Behauptungen ergeben sich unmittelbar aus dieser Definition.

1. Ein lineares Funktional f trennt die Mengen M und N genau dann, wenn es für jedes $x \in L$ die Mengen $M - N$ und $\{0\}$ (d. h. die Menge aller Elemente der Form $x - y$, $x \in M$, $y \in N$, und den Punkt 0) trennt.

2. Ein lineares Funktional f trennt die Mengen M und N genau dann, wenn es die Mengen $M - x$ und $N - x$ für jedes $x \in L$ trennt.

Aus dem Satz von Hahn-Banach erhält man leicht den folgenden Satz über die Trennbarkeit konvexer Mengen in einem linearen Raum, der zahlreiche Anwendungen besitzt.

Satz 5. *Es seien M und N disjunkte konvexe Mengen in einem reellen linearen Raum L, von denen wenigstens eine, etwa M, einen nichtleeren offenen Kern besitzen möge (d. h. ein konvexer Körper ist). Dann existiert ein nicht identisch verschwindendes lineares Funktional f auf L, das M und N trennt.*

Beweis. Ohne Beschränkung der Allgemeinheit kann man annehmen, daß der Punkt 0 zum offenen Kern der Menge M gehört. (Sonst betrachten wir die Mengen $M - x_0$ und $N - x_0$, wobei x_0 irgendein Punkt aus $J(M)$ ist.) Wenn y_0 ein Punkt der Menge N ist, gehört der Punkt $-y_0$ zum offenen Kern der Menge $M - N$, und 0 gehört zum offenen Kern der Menge $K = M - N + y_0$. Da M und N disjunkt sein sollen, ist $0 \notin M - N$ und $y_0 \notin K$. Ist p das Minkowskische Funktional für die Menge K, so ist $p(y_0) \geqq 1$ (wegen $y_0 \notin K$). Wir führen nun das lineare Funktional

$$f_0(\alpha y_0) = \alpha p(y_0)$$

ein. Es ist auf dem eindimensionalen Teilraum aller Elemente der Form αy_0 definiert und genügt wegen $p(\alpha y_0) = \alpha p(y_0)$ für $\alpha \geqq 0$ und $f_0(\alpha y_0) = \alpha f_0(y_0) < 0 < p(\alpha y_0)$ für $\alpha < 0$ der Bedingung

$$f_0(\alpha y_0) \leqq p(\alpha y_0).$$

Nach dem Satz von Hahn-Banach kann man das Funktional f_0 zu einem linearen Funktional f fortsetzen, das auf ganz L definiert ist und auf L der Bedingung $f(y) \leqq p(y)$ genügt. Hieraus ergibt sich $f(y) \leqq 1$ für $y \in K$ und gleichzeitig $f(y_0) \geqq 1$. Somit trennt f die Mengen K und $\{y_0\}$ und folglich auch die Mengen $M - N$ und $\{0\}$. Dann trennt aber f die Mengen M und N. Damit ist der Satz bewiesen.

3.3. Normierte Räume

In Kapitel 2 haben wir uns mit topologischen und insbesondere mit metrischen Räumen beschäftigt, d. h. mit Mengen, für deren Elemente auf irgendeine Weise ein Umgebungsbegriff erklärt war. In den vorangegangenen Abschnitten dieses Kapitels hatten wir es mit linearen Räumen zu tun. Bis jetzt stand jeder dieser Begriffe für

sich allein. Jedoch hat man es in der Analysis oft mit Räumen zu tun, in denen sowohl eine Addition von Elementen und eine Multiplikation mit Zahlen als auch eine Topologie erklärt sind, d. h., man muß sogenannte *topologische lineare Räume* betrachten. Unter diesen bilden die *normierten Räume* eine wichtige Klasse. Die Theorie dieser Räume ist in den Arbeiten S. BANACHS und einer Reihe anderer Mathematiker entwickelt worden.

3.3.1. Definition und Beispiele normierter Räume

Definition 1. L sei ein linearer Raum. Ein endliches konvexes Funktional p auf L heißt *Norm*, wenn (außer der Konvexität) folgende Bedingungen erfüllt sind:

1. $p(x) = 0$ gilt nur für $x = 0$,
2. $p(\alpha x) = |\alpha|\, p(x)$ für alle α.

Somit können wir unter Berücksichtigung der Konvexitätsdefinition sagen, daß jedes endliche Funktional in L eine Norm ist, wenn es den folgenden drei Bedingungen genügt:

1. $p(x) \geqq 0$, wobei $p(x) = 0$ nur für $x = 0$ gilt,
2. $p(x + y) \leqq p(x) + p(y)$, $x, y \in L$,
3. $p(\alpha x) = |\alpha|\, p(x)$ für jede Zahl α.

Definition 2. Einen linearen Raum L, in dem eine Norm erklärt ist, nennen wir *normierten Raum*. Die Norm eines Elementes $x \in L$ bezeichnen wir mit $\|x\|$.

Jeder normierte Raum wird zu einem metrischen Raum, wenn man darin den Abstand

$$\varrho(x, y) = \|x - y\|$$

einführt. Die Gültigkeit der Axiome eines metrischen Raumes folgt sofort aus den Eigenschaften 1 bis 3 einer Norm. Somit lassen sich alle die Begriffe und Ergebnisse, die in Kapitel 2 für metrische Räume angegeben wurden, auf normierte Räume übertragen.

Ein vollständiger normierter Raum heißt *Banachraum* oder kürzer *B-Raum*.

Beispiele normierter Räume. Viele der Räume, die in Kapitel 2 als Beispiele metrischer Räume angegeben wurden, können in natürlicher Weise mit der Struktur eines normierten Raumes versehen werden.

1. Die reellen Zahlen $\boldsymbol{R}^1$ werden zu einem normierten Raum, wenn man $\|x\| = |x|$ für jede Zahl $x \in \boldsymbol{R}^1$ setzt.

2. Setzt man im reellen n-dimensionalen Raum $\boldsymbol{R}^n$

$$\|x\| = \sqrt{\sum_{k=1}^{n} x_k^2}\,, \tag{1}$$

für alle Elemente

$$x = (x_1, x_2, \ldots, x_n),$$

so sind alle Normaxiome erfüllt. Die Formel

$$\varrho(x, y) = \|x - y\| = \sqrt{\sum_{k=1}^{n} (x_k - y_k)^2}$$

definiert in $\boldsymbol{R}^n$ dieselbe Metrik, die wir in diesem Raum bereits betrachtet haben.

In demselben linearen Raum kann man auch die Norm

$$\|x\|_1 = \sum_{k=1}^{n} |x_k| \tag{2}$$

oder die Norm

$$\|x\|_0 = \max_{1 \leq k \leq n} |x_k| \tag{3}$$

einführen. Diese Normen definieren in $\boldsymbol{R}^n$ die Metriken, die wir in 2.1.1., Beispiel 4 und 5, betrachtet haben. Der Nachweis, daß in jedem dieser Fälle tatsächlich die Normaxiome erfüllt sind, bereitet keine Mühe.

Im komplexen n-dimensionalen Raum $\boldsymbol{C}^n$ kann man die Norm

$$\|x\| = \sqrt{\sum_{k=1}^{n} |x_k|^2}$$

oder eine der Normen (2) oder (3) einführen.

3. Im Raum $C[a, b]$ aller stetigen Funktionen auf dem Intervall $[a, b]$ wird durch die Formel

$$\|f\| = \max_{a \leq t \leq b} |f(t)| \tag{4}$$

eine Norm definiert. Der zugehörige Abstand wurde schon in 2.1.1., Beispiel 6, betrachtet.

4. Im Raum m aller beschränkten Zahlenfolgen

$$x = (x_1, x_2, \ldots, x_n, \ldots)$$

setzen wir

$$\|x\| = \sup_n |x_n|. \tag{5}$$

Dann sind die Axiome 1 bis 3 der Normdefinition offensichtlich erfüllt. Die Metrik, die in m durch diese Norm erzeugt wird, stimmt mit der bereits in 2.1.1., Beispiel 9, betrachteten Metrik überein.

3.3.2. Teilräume normierter Räume. Eine nichtleere Menge L_0 eines linearen Raumes L (ohne Topologie) haben wir Teilraum von L genannt, wenn mit $x, y \in L_0$ auch $\alpha x + \beta y \in L_0$ ist. In normierten Räumen sind die *abgeschlossenen* linearen Teilräume von grundlegendem Interesse, d. h. die Teilräume, die alle ihre Häufungspunkte enthalten. In einem endlichdimensionalen normierten Raum ist jeder Teilraum automatisch abgeschlossen. (Man beweise das!) Im unendlichdimensionalen Fall ist das jedoch nicht so. Zum Beispiel bilden die Polynome in Raum $C[a, b]$ der stetigen Funktionen mit der Norm (4) einen Teilraum, der nicht abgeschlossen ist.[1])

Als weiteres Beispiel betrachten wir im Raum m aller beschränkten Folgen die Folgen mit nur endlich vielen von Null verschiedenen Folgengliedern. Diese bilden einen Teilraum von m, der jedoch bezüglich der Norm (5) nicht abgeschlossen ist. In seinem Abschluß ist nämlich zum Beispiel die Folge $(1, 1/2, \ldots, 1/n, \ldots)$ enthalten.

In der Regel werden wir nur abgeschlossene Teilräume betrachten, deshalb ist es zweckmäßig, die in 3.1. eingeführte Terminologie abzuändern. *Teilraum* eines normierten Raumes werden wir jetzt nur noch einen *abgeschlossenen* Teilraum nennen, insbesondere werden wir also jetzt den kleinsten *abgeschlossenen* Teilraum, der ein gegebenes System $\{x_\alpha\}$ von Elementen enthält, als den von $\{x_\alpha\}$ erzeugten Teilraum bezeichnen. Wir werden ihn auch *linearen Abschluß* des Systems $\{x_\alpha\}$ nennen. Die (nicht abgeschlossene) Gesamtheit der Elemente, die mit x und y auch jede Linearkombination $\alpha x + \beta y$ enthält, werden wir *lineare Menge* nennen.

Ein System von Elementen eines normierten Raumes E nennen wir *vollständig*, wenn der davon erzeugte (abgeschlossene!) Teilraum ganz E ist. Zum Beispiel ist nach dem Weierstraßschen Approximationssatz die Gesamtheit der Funktionen $1, t, t^2, \ldots, t^n, \ldots$ im Raum $C[a, b]$ der stetigen Funktionen vollständig.

Aufgaben

1. Es sei R ein Banachraum und $B_1 \supset B_2 \supset \cdots \supset B_n \supset \cdots$ eine Folge von ineinandergeschalteten abgeschlossenen Kugeln in R. Man zeige, daß diese Folge einen nichtleeren Durchschnitt besitzt (ohne vorauszusetzen, daß die Radien dieser Kugeln gegen 0 streben; vgl. Aufgabe 3 auf S. 74). Man gebe ein Beispiel einer Folge von ineinandergeschachtelten nichtleeren beschränkten abgeschlossenen konvexen Mengen in einem B-Raum an, die einen leeren Durchschnitt besitzt.

2. Es sei R ein unendlichdimensionaler B-Raum; dann ist die algebraische Dimension von R (vgl. Aufgabe 3 auf S. 127) überabzählbar.

3. Es sei R ein Banachraum und M ein abgeschlossener Teilraum von R. Wir betrachten den Faktorraum $P = R/M$ und definieren darin eine Norm, indem wir für jede Nebenklasse ξ

$$\|\xi\| = \inf_{x \in \xi} \|x\|$$

setzen. Es ist zu beweisen, daß das auf diese Weise definierte Funktional tatsächlich eine Norm in P ist und daß P vollständig bezüglich dieser Norm ist.

---

[1]) Nach dem Weierstraßschen Approximationssatz ist jede auf einem abgeschlossenen Intervall stetige Funktion Grenzwert einer gleichmäßig konvergenten Folge von Polynomen, so daß der Abschluß des Teilraums der Polynome in $C[a, b]$ ganz $C[a, b]$ ergibt.

4. Es sei R ein linearer normierter Raum. Man zeige die Richtigkeit der folgenden Aussagen:

a) Jeder endlichdimensionale lineare Teilraum in R ist abgeschlossen.

b) Ist M ein abgeschlossener und N ein endlichdimensionaler Teilraum von R, dann ist ihre Summe

$$M + N = \{x : x = y + z,\ y \in M, z \in N\}$$

ein abgeschlossener linearer Raum. Man gebe ein Beispiel zweier abgeschlossener linearer Teilräume in l_2 an, deren Summe nicht abgeschlossen ist.

c) Es sei Q eine offene konvexe Menge in R und $x_0 \notin Q$. Dann existiert eine abgeschlossene Hyperebene, die durch den Punkt x_0 hindurchgeht und Q nicht schneidet.

5. Zwei Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ im linearen Raum R heißen *äquivalent*, wenn Konstanten a, $b > 0$ existieren, so daß $a\,\|x\|_1 \leqq \|x\|_2 \leqq b\,\|x\|_1$ für alle $x \in R$ gilt. Es ist zu beweisen, daß in einem endlichdimensionalen Raum R je zwei Normen äquivalent sind.

3.4. Unitäre Räume

3.4.1. Definition unitärer Räume. In vielen Fällen kann man die Norm in einem linearen Raum mit Hilfe eines Skalarproduktes einführen. Dabei verstehen wir unter einem *Skalarprodukt* in einem reellen linearen Raum R eine reellwertige Funktion (x, y), die für jedes Paar $x, y \in R$ definiert ist und die folgende Eigenschaften besitzt:

1. $(x, y) = (y, x)$,
2. $(x_1 + x_2, y) = (x_1, y) + (x_2, y)$,
3. $(\lambda x, y) = \lambda(x, y)$,
4. $(x, x) \geqq 0$, wobei $(x, x) = 0$ nur für $x = 0$ gilt.

Ein linearer Raum mit einem darin fixierten Skalarprodukt heißt *unitärer Raum*. Die Norm in einem unitären Raum wird durch die Formel

$$\|x\| = \sqrt{(x, x)}$$

definiert. Aus den Eigenschaften 1 bis 4 eines Skalarproduktes folgt, daß dann alle Normaxiome erfüllt sind.

Die Gültigkeit der Axiome 1 und 3 (vgl. 3.3.1.) ist offensichtlich, und die Gültigkeit von Axiom 2 (Dreiecksungleichung) folgt aus der *Schwarzschen Ungleichung*

$$|(x, y)| \leqq \|x\| \cdot \|y\|, \tag{1}$$

die wir jetzt beweisen werden.

Dazu betrachten wir den quadratischen Ausdruck in der reellen Veränderlichen λ

$$\varphi(\lambda) = (\lambda x + y, \lambda x + y) = \lambda^2(x, x) + 2\lambda(x, y) + (y, y)$$
$$= \|x\|^2 \lambda^2 + 2(x, y)\,\lambda + \|y\|^2.$$

Nach Eigenschaft 4 eines Skalarproduktes gilt dann $\varphi(\lambda) \geqq 0$. Also ist die Diskriminante des quadratischen Ausdrucks $\varphi(\lambda)$ kleiner oder gleich Null. Die Schwarzsche Ungleichung (1) ist aber gleichwertig mit der Nichtpositivität der Diskriminante des quadratischen Ausdrucks $\varphi(\lambda)$.

Wir vermerken, daß in einem unitären Raum die Summe, die Skalarmultiplikation und das Skalarprodukt stetig sind, d. h., aus $x_n \to x$, $y_n \to y$ (im Sinne der Normkonvergenz), $\lambda_n \to \lambda$ (als Zahlenfolge) ergibt sich

$$x_n + y_n \to x + y,$$

$$\lambda_n x_n \to \lambda x,$$

$$(x_n, y_n) \to (x, y).$$

Der Beweis dieser Beziehungen stützt sich auf die Benutzung der Schwarzschen Ungleichung (1) und wird dem Leser als Übung überlassen.

Das Vorhandensein eines Skalarproduktes in R gestattet es, in diesem Raum nicht nur die Norm (d. h. die Länge) eines Vektors einzuführen, sondern auch den Winkel zwischen Vektoren. Und zwar wird der Winkel φ zwischen den Vektoren x und y durch die Formel

$$\cos \varphi = \frac{(x, y)}{\|x\| \cdot \|y\|} \tag{2}$$

definiert. Dabei folgt aus der Schwarzschen Ungleichung, daß der Ausdruck auf der rechten Seite von (2) betragsmäßig nicht größer als 1 ist, so daß die Formel (2) tatsächlich für beliebige von Null verschiedene Vektoren x und y einen Winkel φ, $0 \leqq \varphi \leqq \pi$, definiert.

Wenn $(x, y) = 0$ ist, ergibt sich $\varphi = \pi/2$ aus (2); in diesem Fall heißen die Vektoren x und y *orthogonal.*

Ein System von Null verschiedener Vektoren $\{x_\alpha\}$ aus R heißt *orthogonal,* wenn

$$(x_\alpha, x_\beta) = 0 \quad \text{für} \quad \alpha \neq \beta$$

ist. Sind die Vektoren $\{x_\alpha\}$ orthogonal, so sind sie auch linear unabhängig. Denn wenn

$$a_1 x_{\alpha_1} + a_2 x_{\alpha_2} + \cdots + a_n x_{\alpha_n} = 0$$

ist, erhalten wir auf Grund der Orthogonalität des Systems $\{x_\alpha\}$

$$(x_{\alpha_i}, a_1 x_{\alpha_1} + \cdots + a_n x_{\alpha_n}) = a_i (x_{\alpha_i}, x_{\alpha_i}) = 0,$$

und wegen $(x_{\alpha_i}, x_{\alpha_i}) \neq 0$ bedeutet das $a_i = 0$ für alle $i = 1, 2, \ldots, n$.

Wenn ein orthogonales System $\{x_\alpha\}$ *vollständig* ist (d. h., wenn der davon erzeugte abgeschlossene Teilraum ganz R ist), wird es *orthogonale Basis* genannt. Ist dabei die Norm eines jeden Elementes gleich 1, so heißt das System $\{x_\alpha\}$ *orthogonale normierte*

Basis oder kürzer *orthonormierte Basis*. Allgemein heißt ein (vollständiges oder nichtvollständiges) System $\{x_\alpha\}$ mit

$$(x_\alpha, x_\beta) = \begin{cases} 0 & \text{für} \quad \alpha \neq \beta, \\ 1 & \text{für} \quad \alpha = \beta \end{cases}$$

orthogonales normiertes oder kürzer *orthonormiertes System*. Offenbar ist $\left\{\frac{x_\alpha}{\|x_\alpha\|}\right\}$ ein orthonormiertes System, wenn $\{x_\alpha\}$ ein orthogonales System war.

3.4.2. Beispiele. Wir betrachten einige Beispiele unitärer Räume und orthogonaler Basen.

1. Der n-dimensionale euklidische Raum $\boldsymbol{R}^n$ aller n-Tupel

$$x = (x_1, x_2, \ldots, x_n)$$

reeller Zahlen mit der üblichen Addition und Skalarmultiplikation und dem Skalarprodukt

$$(x, y) = \sum_{i=1}^{n} x_i y_i \tag{3}$$

stellt ein wohlbekanntes Beispiel eines unitären Raumes dar. Die Vektoren

$$\begin{aligned} e_1 &= (1, 0, 0, \ldots, 0), \\ e_2 &= (0, 1, 0, \ldots, 0), \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ e_n &= (0, 0, 0, \ldots, 1) \end{aligned}$$

bilden in diesem Raum eine orthonormierte Basis (eine von unendlich vielen möglichen).

2. Der Raum l_2 aller Elemente

$$x = (x_1, x_2, \ldots, x_n, \ldots) \quad \text{mit} \quad \sum_{i=1}^{\infty} x_i^2 < \infty$$

bildet mit dem Skalarprodukt

$$(x, y) = \sum_{i=1}^{\infty} x_i y_i \tag{4}$$

einen unitären Raum. Die Konvergenz der Reihe in (4) wurde schon in 2.1. gezeigt. Die Eigenschaften 1 bis 4 eines Skalarproduktes lassen sich unmittelbar nachprüfen.

Die Vektoren

$$\left.\begin{aligned} e_1 &= (1, 0, 0, \ldots), \\ e_2 &= (0, 1, 0, \ldots), \\ e_3 &= (0, 0, 1, \ldots), \\ &\ldots\ldots\ldots\ldots\ldots \end{aligned}\right\} \tag{5}$$

bilden die einfachste orthonormierte Basis in l_2. Die Orthogonalität und Normiertheit dieses Systems sind klar, außerdem ist das System (5) vollständig. Denn für einen beliebigen Vektor $x = (x_1, x_2, \ldots, x_n, \ldots)$ aus l_2 ist $x^{(n)} = (x_1, x_2, \ldots, x_n, 0, 0, \ldots)$ eine Linearkombination der Vektoren $e_1, \ldots, e_n$, und es gilt

$$\|x^{(n)} - x\| \to 0 \quad \text{für} \quad n \to \infty.$$

3. Der Raum $C^2[a, b]$ aller auf dem Intervall $[a, b]$ stetigen reellen Funktionen mit dem Skalarprodukt

$$(f, g) = \int_a^b f(t)\, g(t)\, dt \tag{6}$$

bildet ebenfalls einen unitären Raum. Unter den verschiedenen orthogonalen Basen, die man in diesem Raum angeben kann, ist das trigonometrische System

$$\frac{1}{2}, \cos n\,\frac{2\pi t}{b-a}, \sin n\,\frac{2\pi t}{b-a} \qquad (n = 1, 2, \ldots) \tag{7}$$

am wichtigsten. Die Orthogonalität dieses Systems läßt sich unmittelbar nachprüfen.

Werden die stetigen Funktionen auf einem Intervall der Länge 2π, z. B. auf $[-\pi, \pi]$, betrachtet, so lautet das entsprechende trigonometrische System

$$1/2, \cos nt, \sin nt \qquad (n = 1, 2, \ldots).$$

Das System (7) ist vollständig. Nach dem Weierstraßschen Approximationssatz kann nämlich jede auf dem Intervall $[a, b]$ stetige Funktion φ mit $\varphi(a) = \varphi(b)$ als Grenzwert einer gleichmäßig konvergenten Folge von trigonometrischen Polynomen, d. h. von Linearkombinationen der Elemente des Systems (7), dargestellt werden.

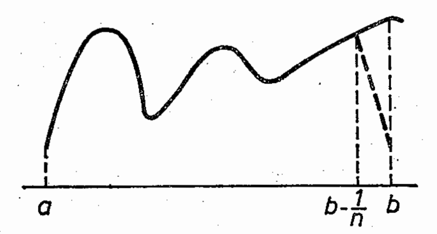


Abb. 17

Eine solche Folge konvergiert aber erst recht in der Norm des Raumes $C^2[a, b]$ gegen φ, und wenn nun f eine beliebige Funktion aus $C^2[a, b]$ ist, kann man sie (in der Norm des Raumes $C^2[a, b]$) als Grenzwert der Folge φ_n, $\varphi_n(x) = f(x)$ für $x \in [a, b - 1/n]$, $\varphi(b) = f(a)$, φ linear auf $[b - 1/n, b]$, darstellen, die im Punkt b denselben Wert wie im Punkt a annimmt (Abb. 17). Daher kann man jedes Element aus $C^2[a, b]$ beliebig gut (in der Metrik dieses Raumes) durch Linearkombinationen von Elementen des Systems (7) annähern, und das bedeutet gerade die Vollständigkeit dieses Systems.

3.4.3. Die Existenz orthogonaler Basen. Orthogonalisierung. Im verbleibenden Teil von 3.4. werden wir uns auf separable unitäre Räume beschränken (die also eine abzählbare überall dichte Menge enthalten). Jeder der Räume aus dem vorangegangenen Abschnitt ist separabel. (Man beweise das!) Ein Beispiel eines nichtseparablen unitären Raumes kann man folgendermaßen konstruieren. Wir betrachten auf der reellen Achse alle die Funktionen x, für die die Menge $T_x = \{t : x(t) \neq 0\}$ höchstens abzählbar ist und für die $\sum_{t \in T_x} x^2(t) < \infty$ ist. Die Addition und die Skalarmultiplikation definieren wir in diesem Raum wie die gewöhnliche Addition und Skalarmultiplikation von Funktionen, und das Skalarprodukt definieren wir durch die Formel

$$(x, y) = \sum_{t \in T_x \cap T_y} x(t)\, y(t).$$

Den Beweis, daß in diesem Raum keine abzählbare überall dichte Teilmenge existiert, überlassen wir dem Leser. Wir bemerken noch, daß dieser Raum vollständig ist.

Es sei also R ein separabler unitärer Raum. Wir werden zeigen, daß *in einem solchen Raum jedes orthogonale System höchstens abzählbar ist.*

Dazu kann man ohne Beschränkung der Allgemeinheit annehmen, daß das betrachtete System $\{\varphi_\alpha\}$ nicht nur orthogonal, sondern auch normiert ist (sonst ersetzen wir es durch das System $\left\{\frac{\varphi_\alpha}{\|\varphi_\alpha\|}\right\}$). Dann gilt

$$\|\varphi_\alpha - \varphi_\beta\| = \sqrt{2} \quad \text{für} \quad \alpha \neq \beta.$$

Wir betrachten jetzt die Gesamtheit der Kugeln $B(\varphi_\alpha, 1/2)$. Diese Kugeln sind paarweise disjunkt. Wenn nun in R eine abzählbare überall dichte Menge $\{\psi_n\}$ existiert, gibt es in jeder Kugel wenigstens ein Element aus $\{\psi_n\}$. Folglich ist die Anzahl dieser Kugeln (und damit auch der Elemente φ_α) höchstens abzählbar.

In jedem der oben angeführten Beispiele unitärer Räume haben wir je eine orthogonale Basis angegeben. Wir werden jetzt einen allgemeinen Satz beweisen, der dem Satz von der Existenz einer orthogonalen Basis im n-dimensionalen euklidischen Raum analog ist.

Satz 1 (*Schmidtsches Orthogonalisierungsverfahren*). *Es sei*

$$f_1, f_2, \ldots, f_n, \ldots \tag{8}$$

ein linear unabhängiges System von Elementen in einem unitären Raum R. Dann existiert in R ein System von Elementen

$$\varphi_1, \varphi_2, \ldots, \varphi_n, \ldots, \tag{9}$$

das den folgenden Bedingungen genügt:

1. *Das System* (9) *ist orthonormiert.*

2. *Jedes Element* φ_n *ist eine Linearkombination der Elemente* $f_1, f_2, \ldots, f_n$,

$$\varphi_n = a_{n1} f_1 + \cdots + a_{nn} f_n, \qquad a_{nn} \neq 0.$$

3. *Jedes Element* f_n *läßt sich in der Form*

$$f_n = b_{n1}\varphi_1 + \cdots + b_{nn}\varphi_n \quad \textit{mit} \quad b_{nn} \neq 0$$

darstellen.

Jedes Element des Systems (9) *wird durch die Bedingungen* 1 *bis* 3 *eindeutig bis auf den Faktor* ± 1 *bestimmt.*

Beweis. Wir setzen

$$\varphi_1 = a_{11} f_1,$$

dabei bestimmen wir a_{11} aus der Bedingung

$$(\varphi_1, \varphi_1) = a_{11}^2 (f_1, f_1) = 1,$$

woraus

$$a_{11} = \frac{1}{b_{11}} = \frac{\pm 1}{\sqrt{(f_1, f_1)}}$$

folgt. Es ist klar, daß φ_1 dadurch eindeutig (bis aufs Vorzeichen) bestimmt ist. Es seien die Elemente φ_k $(k < n)$ bereits konstruiert, so daß die Bedingungen 1 bis 3 erfüllt sind. Dann kann man f_n in der Form

$$f_n = b_{n1}\varphi_1 + \cdots + b_{n,n-1}\varphi_{n-1} + h_n$$

mit

$$(h_n, \varphi_k) = 0 \quad \text{für} \quad k < n$$

darstellen. Die Koeffizienten b_{nk} und damit auch das Element h_n werden eindeutig bestimmt durch die Bedingungen

$$\begin{aligned}(h_n, \varphi_k) &= (f_n - b_{n1}\varphi_1 - \cdots - b_{n,n-1}\varphi_{n-1}, \varphi_k)\\ &= (f_n, \varphi_k) - b_{nk}(\varphi_k, \varphi_k) = 0.\end{aligned}$$

Offensichtlich gilt dabei $(h_n, h_n) > 0$ (die Annahme $(h_n, h_n) = 0$ würde der linearen Unabhängigkeit des Systems (8) widersprechen). Wir setzen dann

$$\varphi_n = \frac{h_n}{\sqrt{(h_n, h_n)}}.$$

Aus der induktiven Konstruktion geht hervor, daß sich h_n und damit auch φ_n durch $f_1, \ldots, f_n$ ausdrücken lassen, d. h., es ist

$$\varphi_n = a_{n1}f_1 + \cdots + a_{nn}f_n \quad \text{mit} \quad a_{nn} = \frac{1}{\sqrt{(h_n, h_n)}} \neq 0 .$$

Außerdem gilt

$$(\varphi_n, \varphi_n) = 1, \quad (\varphi_n, \varphi_k) = 0 \qquad (k < n)$$

und

$$f_n = b_{n1}\varphi_1 + \cdots + b_{nn}\varphi_n \qquad \left(b_{nn} = \sqrt{(h_n, h_n)} \neq 0\right),$$

d. h., φ_n erfüllt die Bedingungen des Satzes.

Der Übergang vom System (8) zum System (9), das den Bedingungen 1 bis 3 genügt, heißt *Schmidtsches Orthogonalisierungsverfahren.*

Es ist klar, daß die von den Systemen (8) und (9) erzeugten Teilräume übereinstimmen. Somit sind diese Systeme entweder gleichzeitig vollständig oder gleichzeitig nicht vollständig.

Folgerung. *In einem separablen unitären Raum R existieren orthonormierte Basen.*

Zum Beweis sei $\psi_1, \psi_2, \ldots, \psi_n, \ldots$ eine abzählbare überall dichte Menge in R. In dieser Menge kann man ein vollständiges System $\{f_n\}$ linear unabhängiger Elemente auswählen. Dazu braucht man nur in der Folge $\{\psi_n\}$ alle die Elemente ψ_k wegzulassen, die als Linearkombination der Elemente ψ_i mit $i < k$ dargestellt werden können. Wenn wir auf das somit erhaltene vollständige System von linear unabhängigen Vektoren das Orthogonalisierungsverfahren anwenden, erhalten wir eine orthonormierte Basis.

Aufgaben

1. Es ist ein Beispiel eines (nichtseparablen) unitären Raumes anzuführen, in dem keine orthogonale Basis existiert. Man beweise, daß in einem vollständigen (nicht notwendig separablen) unitären Raum eine orthonormierte Basis existiert.

2. Man beweise, daß in einem vollständigen (nicht notwendig separablen) unitären Raum jede Folge von ineinandergeschachtelten nichtleeren beschränkten konvexen abgeschlossenen Mengen einen nichtleeren Durchschnitt besitzt (vgl. die Aufgaben auf S. 74 und S. 144).

3.4.4. Die Besselsche Ungleichung. Abgeschlossene orthogonale Systeme. Wenn man im n-dimensionalen euklidischen Raum eine orthonormierte Basis $e_1, e_2, \ldots, e_n$ auswählt, kann man jeden Vektor $x \in \boldsymbol{R}^n$ in der Form

$$x = \sum_{k=1}^{n} c_k e_k \tag{10}$$

mit

$$c_k = (x, e_k) \tag{11}$$

darstellen. Wir wollen klären, wie die Darstellung (10) auf den Fall des unendlichdimensionalen euklidischen Raumes zu verallgemeinern ist. Dazu sei

$$\varphi_1, \varphi_2, \ldots, \varphi_n, \ldots \tag{12}$$

ein orthonormiertes System im unitären Raum R und f ein beliebiges Element aus R. Wir ordnen dem Element $f \in R$ die Folge der Zahlen

$$c_k = (f, \varphi_k), \qquad k = 1, 2, \ldots, \tag{13}$$

die wir *Koordinaten* oder *Fourierkoeffizienten* von f bezüglich des Systems $\{\varphi_k\}$ nennen, und die (vorläufig formale) Reihe

$$\sum_k c_k \varphi_k \tag{14}$$

zu, die wir *Fourierreihe* des Elements f bezüglich des Systems $\{\varphi_n\}$ nennen.

Natürlich erhebt sich die Frage, ob die Reihe (14) konvergiert, d. h., ob die Folge ihrer Teilsummen (im Sinne der Metrik von R) gegen einen Grenzwert strebt, und ob dieser Grenzwert im Fall der Konvergenz mit dem Ausgangselement f übereinstimmt.

Um diese Fragen zu beantworten, betrachten wir zunächst folgende Aufgabe. Bei gegebenem n sind die Koeffizienten α_k $(k = 1, 2, \ldots, n)$ so zu bestimmen, daß der Abstand zwischen f und der Summe

$$S_n = \sum_{k=1}^{n} \alpha_k \varphi_k \tag{15}$$

minimal wird. Dazu berechnen wir diesen Abstand. Da das System (12) orthonormiert ist, gilt

$$\begin{aligned}
\|f - S_n\|^2 &= \left(f - \sum_{k=1}^{n} \alpha_k \varphi_k, f - \sum_{k=1}^{n} \alpha_k \varphi_k\right) \\
&= (f, f) - 2\left(f, \sum_{k=1}^{n} \alpha_k \varphi_k\right) + \left(\sum_{k=1}^{n} \alpha_k \varphi_k, \sum_{k=1}^{n} \alpha_k \varphi_k\right) \\
&= \|f\|^2 - 2\sum_{k=1}^{n} \alpha_k c_k + \sum_{k=1}^{n} \alpha_k^2 \\
&= \|f\|^2 - \sum_{k=1}^{n} c_k^2 + \sum_{k=1}^{n} (\alpha_k - c_k)^2 .
\end{aligned}$$

Offenbar wird das Minimum dieses Ausdrucks dann angenommen, wenn der letzte Summand gleich 0 ist, d. h., wenn

$$\alpha_k = c_k \qquad (k = 1, 2, \ldots, n) \tag{16}$$

ist. In diesem Fall gilt

$$\|f - S_n\|^2 = \|f\|^2 - \sum_{k=1}^{n} c_k^2. \tag{17}$$

Wir haben also gezeigt, daß bei gegebenem n die n-te Partialsumme der Fourierreihe von f von allen Summen der Form (15) am wenigsten von f abweicht. Geometrisch kann man dieses Resultat in der folgenden Weise interpretieren. Das Element

$$f - \sum_{k=1}^{n} \alpha_k \varphi_k$$

ist dann und nur dann zu allen Linearkombinationen der Form

$$\sum_{k=1}^{n} \beta_k \varphi_k$$

orthogonal, d. h. orthogonal zu dem von den Elementen $\varphi_1, \varphi_2, \ldots, \varphi_n$ erzeugten Teilraum, wenn die Bedingung (16) erfüllt ist. (Man weise das nach!) Somit stellt das soeben erhaltene Resultat eine Verallgemeinerung eines bekannten Satzes aus der Elementargeometrie dar: Die Länge des von einem gegebenen Punkt auf eine Gerade oder eine Ebene gefällten Lots ist kleiner als die Länge jeder Verbindungsstrecke dieses Punktes mit der Geraden bzw. der Ebene.

Da stets $\|f - S_n\|^2 \geqq 0$ ist, folgt aus Gleichung (17)

$$\sum_{k=1}^{n} c_k^2 \leqq \|f\|^2.$$

Hier ist n beliebig, und die rechte Seite hängt nicht von n ab. Folglich konvergiert die Reihe $\sum_{k=1}^{\infty} c_k^2$, und es ist

$$\sum_{k=1}^{\infty} c_k^2 \leqq \|f\|^2. \tag{18}$$

Diese Ungleichung heißt *Besselsche Ungleichung*. Geometrisch besagt sie, daß die Summe der Quadrate von Projektionen des Vektors f auf paarweise orthogonale Richtungen das Quadrat der Länge des Vektors f nicht übersteigt.

Wir führen nun folgenden wichtigen Begriff ein.

Definition 1. Das orthonormierte System (12) heißt *abgeschlossen*, wenn für jedes $f \in R$ die Gleichung

$$\sum_{k=1}^{\infty} c_k^2 = \|f\|^2, \tag{19}$$

die sogenannte *Parsevalsche Gleichung*, gültig ist.

Aus der Identität (17) folgt, daß die Abgeschlossenheit des Systems (12) damit gleichwertig ist, daß für jedes $f \in R$ die Partialsummen der Fourierreihe $\sum_{n=1}^{\infty} c_n \varphi_n$ gegen f konvergieren.

Der Begriff der Abgeschlossenheit eines orthonormierten Systems hängt eng mit dem oben eingeführten Begriff der Vollständigkeit eines Systems zusammen.

Satz 2. *Ein orthonormiertes System in einem separablen unitären Raum ist genau dann vollständig, wenn es abgeschlossen ist.*

Beweis. Das System $\{\varphi_n\}$ sei abgeschlossen; dann konvergiert für jedes $f \in R$ die Folge der Partialsummen der Fourierreihe von f gegen f. Das bedeutet, daß die Linearkombinationen der Elemente des Systems $\{\varphi_n\}$ überall dicht sind in R, d. h., das System $\{\varphi_n\}$ ist vollständig. Umgekehrt sei das System $\{\varphi_n\}$ vollständig, d. h., man kann jedes Element $f \in R$ beliebig genau durch eine Linearkombination

$$\sum_{k=1}^{n} \alpha_k \varphi_k$$

von Elementen des Systems $\{\varphi_n\}$ approximieren. Die Partialsumme

$$\sum_{k=1}^{n} c_k \varphi_k$$

der Fourierreihe von f liefert dann aber eine mindestens ebensogute Approximation für f. Folglich konvergiert die Reihe

$$\sum_{k=1}^{\infty} c_k \varphi_k$$

gegen f, und es gilt die Parsevalsche Gleichung.

Im vorhergehenden Abschnitt haben wir die Existenz vollständiger orthonormierter Systeme in einem separablen unitären Raum bewiesen. Da für orthonormierte Systeme die Begriffe der Abgeschlossenheit und der Vollständigkeit übereinstimmen, brauchen wir die Existenz abgeschlossener orthogonaler Systeme in R nicht mehr zu zeigen, und die im vorangegangenen Abschnitt angeführten Beispiele vollständiger orthonormierter Systeme sind zugleich Beispiele abgeschlossener Systeme.

Wir haben bisher stets angenommen, daß die betrachteten orthogonalen Systeme normiert sind. Man kann die Begriffe des Fourierkoeffizienten, der Fourierreihe usw. aber auch für beliebige orthogonale Systeme formulieren. Es sei $\{\varphi_n\}$ ein beliebiges orthogonales System. Dann kann man daraus das normierte System der Elemente $\psi_n = \varphi_n/\|\varphi_n\|$ konstruieren. Für jedes $f \in R$ gilt nun

$$c_n = (f, \psi_n) = \frac{1}{\|\varphi_n\|} (f, \varphi_n)$$

und

$$\sum_{n=1}^{\infty} c_n \psi_n = \sum_{n=1}^{\infty} \frac{c_n}{\|\varphi_n\|} \varphi_n = \sum_{n=1}^{\infty} a_n \varphi_n$$

mit

$$a_n = \frac{c_n}{\|\varphi_n\|} = \frac{(f, \varphi_n)}{\|\varphi_n\|^2}. \tag{20}$$

Die durch Formel (20) definierten Koeffizienten nennen wir *Fourierkoeffizienten* des Elements f bezüglich des orthogonalen (nicht normierten) Systems $\{\varphi_n\}$. Wenn wir in Ungleichung (18) die Beziehung $c_n = a_n \|\varphi_n\|$ aus (20) einsetzen, erhalten wir

$$\sum_{n=1}^{\infty} \|\varphi_n\|^2 a_n^2 \leqq \|f\|^2, \tag{21}$$

die Besselsche Ungleichung für beliebige orthogonale Systeme.

3.4.5. Vollständige unitäre Räume. Der Satz von Riesz-Fischer. Von 3.4.3. an haben wir separable unitäre Räume betrachtet, von jetzt an werden wir außerdem voraussetzen, daß die betrachteten Räume vollständig sind.

Es sei also R ein vollständiger separabler unitärer Raum und $\{\varphi_n\}$ ein orthonormiertes System in R (das nicht notwendig vollständig sein muß). Aus der Besselschen Ungleichung folgt, daß Zahlen $c_1, c_2, \ldots, c_n, \ldots$ nur dann Fourierkoeffizienten eines Elementes $f \in R$ sein können, wenn die Reihe

$$\sum_{n=1}^{\infty} c_n^2$$

konvergiert. Es zeigt sich, daß diese Bedingung in einem vollständigen Raum nicht nur notwendig, sondern auch hinreichend ist, und zwar gilt der folgende Satz.

Satz 3 (Riesz-Fischer). *Es sei $\{\varphi_n\}$ ein beliebiges orthonormiertes System in einem vollständigen unitären Raum R. Ferner seien*

$$c_1, c_2, \ldots, c_n, \ldots$$

Zahlen, für die die Reihe

$$\sum_{k=1}^{\infty} c_k^2 \tag{22}$$

konvergent ist. Dann existiert ein Element $f \in R$ mit

$$c_k = (f, \varphi_k)$$

und

$$\sum_{k=1}^{\infty} c_k^2 = (f, f) = \|f\|^2.$$

Beweis. Wir setzen

$$f_n = \sum_{k=1}^{n} c_k \varphi_k.$$

Dann gilt

$$\|f_{n+p} - f_n\|^2 = \|c_{n+1}\varphi_{n+1} + \cdots + c_{n+p}\varphi_{n+p}\|^2 = \sum_{k=n+1}^{n+p} c_k^2.$$

Da die Reihe (22) konvergiert, folgt hieraus wegen der Vollständigkeit von R die Konvergenz der Folge $\{f_n\}$ gegen ein Element $f \in R$. Es ist nun

$$(f, \varphi_i) = (f_n, \varphi_i) + (f - f_n, \varphi_i); \tag{23}$$

dabei ist der erste Summand rechts für $n \geqq i$ gleich c_i, und der zweite strebt für $n \to \infty$ wegen

$$|(f - f_n, \varphi_i)| \leqq \|f - f_n\| \cdot \|\varphi_i\|$$

gegen Null.

Die linke Seite der Gleichung (23) hängt nicht von n ab; daher erhalten wir durch den Grenzübergang $n \to \infty$

$$(f, \varphi_i) = c_i.$$

Nach Definition von f gilt

$$\|f - f_n\| \to 0 \quad \text{für} \quad n \to \infty.$$

Daraus folgt

$$(f, f) - \sum_{k=1}^{n} c_k^2 = \left(f - \sum_{k=1}^{n} c_k\varphi_k, f - \sum_{k=1}^{n} c_k\varphi_k\right) \to 0$$

für $n \to \infty$, und das bedeutet gerade

$$\sum_{k=1}^{\infty} c_k^2 = (f, f).$$

Damit ist der Satz bewiesen.

Abschließend geben wir noch den folgenden nützlichen Satz an.

Satz 4. *Ein orthonormiertes System $\{\varphi_n\}$ in einem vollständigen separablen unitären Raum ist dann und nur dann vollständig, wenn in R kein von Null verschiedenes Element existiert, das zu allen Elementen des Systems $\{\varphi_n\}$ orthogonal ist.*

Beweis. Das System $\{\varphi_n\}$ sei vollständig und somit auch abgeschlossen. Wenn f orthogonal zu allen Elementen des Systems $\{\varphi_n\}$ ist, sind alle Fourierkoeffizienten von f gleich Null. Dann erhalten wir aus der Parsevalschen Gleichung

$$(f, f) = \sum_{k=1}^{\infty} c_k^2 = 0,$$

d. h. $f = 0$.

Umgekehrt sei das System $\{\varphi_n\}$ nicht vollständig. Dann existiert in R ein Element $g \neq 0$ mit

$$(g, g) > \sum_{k=1}^{\infty} c_k^2,$$

wobei $c_k = (g, \varphi_k)$ ist. Auf Grund des Satzes von RIESZ-FISCHER gibt es ein Element $f \in R$ mit

$$(f, \varphi_k) = c_k \quad \text{und} \quad (f, f) = \sum_{k=1}^{\infty} c_k^2.$$

Das Element $f - g$ ist dann aber orthogonal zu allen φ_i, und aus der Ungleichung

$$(f, f) = \sum_{k=1}^{\infty} c_k^2 < (g, g)$$

folgt $f - g \neq 0$. Damit ist der Satz bewiesen.

Aufgaben

1. Es sei H ein vollständiger unitärer (nicht notwendig separabler) Raum, dann existiert in H ein vollständiges orthonormiertes System $\{\varphi_\alpha\}$ (vgl. Aufgabe 1 auf S. 151). Es ist zu zeigen, daß für jeden Vektor $f \in H$ die Entwicklungen

$$f = \sum_{\alpha} (f, \varphi_\alpha)\varphi_\alpha, \quad \|f\|^2 = \sum_{\alpha} (f, \varphi_\alpha)^2$$

gelten, wobei in den rechts stehenden Summen höchstens abzählbar viele von 0 verschiedene Summanden vorkommen.

2. Ein System $\{\varphi_\alpha\}$ von Vektoren eines unitären Raumes R heißt *total*, wenn in R kein von 0 verschiedener Vektor existiert, der zu allen φ_α orthogonal ist. Satz 4 besagt dann, daß in einem vollständigen unitären Raum die Totalität eines Systems von Vektoren äquivalent ist mit dessen Vollständigkeit. Es ist zu zeigen, daß in nichtvollständigen Räumen totale, aber nichtvollständige Systeme existieren können.

3.4.6. Der Hilbertraum. Der Isomorphiesatz.

Wir setzen die Betrachtung vollständiger unitärer Räume fort. Dabei werden wir uns, wie schon bisher, für unendlichdimensionale Räume interessieren und nicht für endlichdimensionale, die in den Vorlesungen über lineare Algebra ausführlich behandelt werden. Nach wie vor werden wir dabei in der Regel die Existenz einer abzählbaren überall dichten Menge in den betrachteten Räumen voraussetzen.

Definition 2. Ein unendlichdimensionaler vollständiger unitärer Raum heißt *Hilbertraum*[1].[2])

[1]) Nach dem berühmten deutschen Mathematiker D. HILBERT (1862—1943), der diesen Begriff einführte.

[2]) Anm. d. Übers.: In der Literatur werden häufig die vollständigen unitären Räume als Hilberträume bezeichnet. Die obige Definition schließt die endlichdimensionalen unitären Räume aus, die im folgenden auch *euklidische Räume* genannt werden.

Somit heißt eine Gesamtheit H von Elementen $f, g, \ldots$ beliebiger Natur Hilbertraum, wenn sie folgenden Bedingungen (Axiomen) genügt:

I. H ist ein unitärer Raum (d. h. ein linearer Raum mit Skalarprodukt).

II. Der Raum H ist vollständig in der Metrik $\varrho(f, g) = \|f - g\|$.

III. Der Raum H ist unendlichdimensional, d. h., man kann in H für jedes n stets n linear unabhängige Elemente finden.

Meist werden separable Hilberträume betrachtet, d. h. Räume, die noch ein zusätzliches Axiom erfüllen.

IV. H ist separabel, d. h., in H existiert eine abzählbare überall dichte Teilmenge.

Als Beispiel eines separablen Hilbertraumes kann der reelle Raum l_2 dienen.

Im weiteren werden wir nur den separablen Fall betrachten, ohne das besonders zu betonen.

Wir erinnern daran, daß zwei unitäre Räume R und R^* isomorph heißen, wenn man zwischen ihren Elementen eine eineindeutige Beziehung herstellen kann, so daß mit

$$x \leftrightarrow x^*, \quad y \leftrightarrow y^* \qquad (x, y \in R; x^*, y^* \in R^*)$$

auch

$$x + y \leftrightarrow x^* + y^*,$$

$$\alpha x \leftrightarrow \alpha x^*$$

und

$$(x, y) \leftrightarrow (x^*, y^*)$$

gilt. Anders ausgedrückt ist ein Isomorphismus unitärer Räume eine eineindeutige Beziehung, die sowohl die linearen Operationen in diesen Räumen als auch das Skalarprodukt erhält.

Bekanntlich sind zwei beliebige n-dimensionale unitäre Räume zueinander isomorph, und somit ist ein solcher Raum isomorph zum Raum $\boldsymbol{R}^n$ (3.4.2., Beispiel 1). Unendlichdimensionale unitäre Räume sind nicht notwendig zueinander isomorph. Zum Beispiel sind die Räume l_2 und $C^2[a, b]$ nicht isomorph. Das ist etwa daraus ersichtlich, daß l_2 vollständig ist und $C^2[a, b]$ nicht.

Es gilt jedoch der folgende Sachverhalt.

Satz 5. *Zwei beliebige separable Hilberträume sind zueinander isomorph.*

Beweis. Wir werden zeigen, daß jeder Hilbertraum H isomorph zum Raum l_2 ist, womit der Satz bewiesen ist. Zu diesem Zweck wählen wir in H ein beliebiges vollständiges orthonormiertes System $\{\varphi_n\}$ und ordnen jedem Element $f \in H$ die Folge seiner Fourierkoeffizienten $(c_1, c_2, \ldots, c_n, \ldots)$ bezüglich des Systems $\{\varphi_n\}$ zu. Diese Folge ist wegen $\sum_{k=1}^{\infty} c_k^2 < \infty$ ein Element aus l_2. Umgekehrt entspricht nach dem

Satz von RIESZ-FISCHER jedem Element $(c_1, c_2, \ldots, c_n, \ldots)$ aus l_2 ein Element $f \in H$, dessen Fourierkoeffizienten die Zahlen $c_1, c_2, \ldots, c_n, \ldots$ sind. Die so zwischen den Elementen aus H und l_2 hergestellte Beziehung ist eineindeutig. Aus

$$f \leftrightarrow (c_1, c_2, \ldots, c_n, \ldots)$$

und

$$g \leftrightarrow (d_1, d_2, \ldots, d_n, \ldots)$$

folgt dabei

$$f + g \leftrightarrow (c_1 + d_1, c_2 + d_2, \ldots, c_n + d_n, \ldots)$$

und

$$\alpha f \leftrightarrow (\alpha c_1, \alpha c_2, \ldots, \alpha c_n, \ldots),$$

d. h., die Summe zweier Elemente aus H geht in die Summe der zugeordneten Elemente aus l_2 und das Produkt eines Elements aus H mit einer Zahl geht in das Produkt des zugeordneten Elements in l_2 mit derselben Zahl über. Schließlich folgt mit Hilfe der Parsevalschen Gleichung die Beziehung

$$(f, g) = \sum_{n=1}^{\infty} c_n d_n, \tag{24}$$

und zwar aus

$$(f, f) = \sum_{n=1}^{\infty} c_n^2, \qquad (g, g) = \sum_{n=1}^{\infty} d_n^2$$

und

$$\begin{aligned}(f + g, f + g) &= (f, f) + 2(f, g) + (g, g) \\ &= \sum_{n=1}^{\infty} (c_n + d_n)^2 \\ &= \sum_{n=1}^{\infty} c_n^2 + 2 \sum_{n=1}^{\infty} c_n d_n + \sum_{n=1}^{\infty} d_n^2.\end{aligned}$$

Somit ist die zwischen den Elementen der Räume H und l_2 hergestellte Beziehung tatsächlich ein Isomorphismus.

Der eben bewiesene Satz besagt, daß es bis auf Isomorphie n u r e i n e n (separablen) Hilbertraum gibt (d. h., das System der Axiome I bis IV ist vollständig) und daß man den Raum l_2 als eine Realisierung des separablen Hilbertraums ansehen kann, ebenso wie der $\boldsymbol{R}^n$ mit dem Skalarprodukt $\sum_{i=1}^{n} x_i y_i$ eine Realisierung des abstrakten n-dimensionalen unitären Raumes darstellt.

Eine andere Realisierung des (separablen) Hilbertraumes kann man erhalten, indem man den Funktionenraum $C^2[a, b]$ nimmt und dessen Vervollständigung betrachtet. Man kann leicht nachprüfen, daß die Vervollständigung R^* jedes unitären

Raumes R (in dem Sinne, wie wir die Vervollständigung eines metrischen Raumes in 2.3. definiert haben) zu einem linearen unitären Raum wird, wenn man die linearen Operationen und das Skalarprodukt von R stetig auf R^* fortsetzt, d. h., wenn man für Elemente $x, y \in R^*$

$$x + y = \lim_{n\to\infty} (x_n + y_n), \qquad \alpha x = \lim_{n\to\infty} \alpha x_n$$

und

$$(x, y) = \lim_{n\to\infty} (x_n, y_n)$$

setzt, falls $\{x_n\}, \{y_n\} \subset R$ approximierende Folgen für x bzw. y sind. (Die Existenz aller dieser Grenzwerte und ihre Unabhängigkeit von der Wahl der Folgen $\{x_n\}$ und $\{y_n\}$ stellt man leicht fest.) Die Vervollständigung des Raumes $C^2[a, b]$ ist somit ein vollständiger unitärer Raum, der offenbar unendlichdimensional und separabel, also ein Hilbertraum ist. In Kapitel 6 werden wir auf diesen Raum zurückkommen und zeigen, daß man die Elemente, die bei der Vervollständigung von $C^2[a, b]$ zu diesem Raum hinzugefügt werden müssen, auch als Funktionen auffassen kann (und zwar als Funktionen, die im Lebesgueschen Sinne quadratisch integrierbar und nicht mehr stetig sind).

3.4.7. Teilräume. Orthogonales Komplement. Direkte Summe. In Übereinstimmung mit den allgemeinen Definitionen aus 3.3. nennen wir eine Menge L im Hilbertraum H *lineare Menge*, wenn mit $f, g \in L$ auch $\alpha f + \beta g \in L$ ist für beliebige Zahlen α und β. Eine abgeschlossene lineare Menge heißt *Teilraum*. Wir führen einige Beispiele von Teilräumen im Hilbertraum an.

1. Es sei h ein beliebiges Element aus H. Dann bildet die Gesamtheit aller Elemente $f \in H$, die zu h orthogonal sind, einen Teilraum in H.

2. Es sei $H = l_2$, d. h. die Gesamtheit aller Zahlenfolgen $(x_1, x_2, \ldots, x_n, \ldots)$ mit $\sum_k x_k^2 < \infty$. Dann bilden die Elemente mit $x_1 = x_2$ einen Teilraum.

3. Es sei wiederum $H = l_2$. Dann bilden die Elemente $x = (x_1, x_2, \ldots, x_n, \ldots)$ mit $x_n = 0$ für $n = 2, 4, 6, \ldots$ (und x_n beliebig für $n = 1, 3, 5, \ldots$) einen Teilraum.

Der Leser möge nachprüfen, daß die in den Beispielen 1 bis 3 angegebenen Mengen von Vektoren tatsächlich Teilräume sind.

Jeder Teilraum eines Hilbertraumes ist entweder ein endlichdimensionaler unitärer Raum oder ebenfalls wieder ein Hilbertraum. Denn offenbar gelten die Axiome I und II für jeden solchen Teilraum, und die Gültigkeit von Axiom IV ergibt sich aus dem folgenden Lemma.

Lemma. *Jede Teilmenge R' eines separablen metrischen Raumes R ist separabel.*

Beweis. Es sei

$$\xi_1, \xi_2, \ldots, \xi_n, \ldots$$

eine abzählbare überall dichte Menge in R und

$$a_n = \inf_{\eta \in R'} \varrho(\xi_n, \eta).$$

Zu beliebigen natürlichen Zahlen n und m gibt es daher einen Punkt $\eta_{n,m} \in R'$ mit

$$\varrho(\xi_n, \eta_{n,m}) < a_n + \frac{1}{m}.$$

Es sei $\varepsilon > 0$ und $1/m < \varepsilon/3$. Dann existiert zu jedem $\eta \in R'$ eine natürliche Zahl n, so daß

$$\varrho(\xi_n, \eta) < \frac{\varepsilon}{3}$$

und folglich

$$\varrho(\xi_n, \eta_{n,m}) < a_n + \frac{1}{m} < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \frac{2\varepsilon}{3}$$

ist. Daraus folgt aber $\varrho(\eta, \eta_{n,m}) < \varepsilon$, d. h., die höchstens abzählbare Menge $\{\eta_{n,m}\}$ $(n, m = 1, 2, \ldots)$ ist überall dicht in R'.

Teilräume eines Hilbertraumes besitzen einige besondere Eigenschaften (die nicht für Teilräume eines beliebigen normierten Raumes gelten). Diese Eigenschaften hängen mit dem Vorhandensein eines Skalarproduktes und dem darauf beruhenden Begriff der Orthogonalität im Hilbertraum zusammen.

Wenden wir das Orthogonalisierungsverfahren im Hilbertraum auf eine abzählbare überall dichte Menge eines Teilraums an, so erhalten wir folgenden Satz.

Satz 6. *In jedem Teilraum M des Hilbertraumes H gibt es ein orthonormiertes System $\{\varphi_n\}$, dessen linearer Abschluß mit M übereinstimmt.*

Es sei M ein Teilraum von H. Dann bezeichnen wir mit

$$M^\perp = H \ominus M$$

die Menge der Elemente $g \in H$, die zu allen Elementen $f \in M$ orthogonal sind. Wir wollen beweisen, daß $M^\perp$ ebenfalls ein Teilraum von H ist. Offenbar ist $M^\perp$ ein linearer Raum, da aus $(g_1, f) = (g_2, f) = 0$ auch $(\alpha_1 g_1 + \alpha_2 g_2, f) = 0$ folgt. Zum Beweis der Abgeschlossenheit nehmen wir an, daß die Elemente g_n zu $M^\perp$ gehören und gegen g konvergieren. Dann gilt für jedes $f \in M$

$$(g, f) = \lim_{n\to\infty} (g_n, f) = 0,$$

und deshalb gehört g ebenfalls zu $M^\perp$. Der Teilraum $M^\perp$ heißt *orthogonales Komplement* des Teilraumes M.

Aus Satz 6 erhält man leicht den folgenden Satz.

Satz 7. *Ist M ein (abgeschlossener!) linearer Teilraum des Raumes H, so kann jedes Element $f \in H$ auf eindeutige Weise in der Form $f = h + h'$ mit $h \in M$ und $h' \in M^\perp$ dargestellt werden.*

Beweis. Wir zeigen zunächst die Existenz einer solchen Darstellung. Dazu wählen wir in M ein vollständiges orthonormiertes System $\{\varphi_n\}$ und setzen

$$h = \sum_{n=1}^{\infty} c_n \varphi_n, \qquad c_n = (f, \varphi_n).$$

Da die Reihe $\sum_{n=1}^{\infty} c_n^2$ (nach der Besselschen Ungleichung) konvergiert, ist h sinnvoll definiert und gehört zu M. Nun setzen wir

$$h' = f - h.$$

Offensichtlich ist dann für alle n

$$(h', \varphi_n) = 0,$$

und weil jedes Element ζ aus M in der Form

$$\zeta = \sum_{n=1}^{\infty} a_n \varphi_n$$

darstellbar ist, erhalten wir

$$(h', \zeta) = \sum_{n=1}^{\infty} a_n (h', \varphi_n) = 0,$$

d. h. $h' \in M^\perp$.

Wir nehmen jetzt an, daß außer der soeben konstruierten Darstellung $f = h + h'$ noch eine andere Darstellung

$$f = h_1 + h_1', \qquad h_1 \in M, \qquad h_1' \in M^\perp,$$

existiert. Dann gilt für alle n

$$(h_1, \varphi_n) = (f, \varphi_n) = c_n,$$

woraus

$$h_1 = h, \qquad h_1' = h'$$

folgt. Damit ist der Satz bewiesen.

Aus Satz 7 ergeben sich einige nützliche Folgerungen.

Folgerung 1. *Das orthogonale Komplement des orthogonalen Komplements eines abgeschlossenen linearen Teilraumes M stimmt mit M überein.*

Somit kann man von zueinander komplementären Teilräumen des Raumes H sprechen. Wenn M und $M^\perp$ zwei solche sich komplementierende Räume und $\{\varphi_n\}$, $\{\varphi_n'\}$ vollständige orthonormierte Systeme in M bzw. $M^\perp$ sind, so ergibt die Vereinigung der Systeme $\{\varphi_n\}$ und $\{\varphi_n'\}$ ein vollständiges orthonormiertes System im ganzen Raum H. Daher gilt:

Folgerung 2. *Jedes orthonormierte System kann zu einem in H vollständigen orthonormierten System erweitert werden.*

Ist das System $\{\varphi_n\}$ endlich, so ist die Anzahl seiner Elemente gleich der Dimension des von $\{\varphi_n\}$ erzeugten Teilraumes M und gleich der Codimension des Teilraumes $M^\perp$. Somit erhalten wir eine weitere Folgerung:

Folgerung 3. *Das orthogonale Komplement eines n-dimensionalen Teilraumes besitzt die Codimension n und umgekehrt.*

Wenn jeder Vektor $f \in H$ in der Form $f = h + h'$, $h \in M$, $h' \in M^\perp$ ($M^\perp$ ist das orthogonale Komplement von M), darstellbar ist, sagt man, daß H die *direkte Summe* der zueinander orthogonalen Teilräume M und $M^\perp$ ist, und schreibt

$$H = M \oplus M^\perp.$$

Es ist klar, daß der Begriff der direkten Summe sofort auf endlich oder sogar abzählbar viele Teilräume verallgemeinert werden kann. Und zwar sagt man, daß H die direkte Summe seiner Teilräume $M_1, M_2, \ldots, M_n, \ldots$ ist,

$$H = M_1 \oplus M_2 \oplus \cdots \oplus M_n \oplus \cdots,$$

wenn

1. die Teilräume M_i paarweise orthogonal sind, d. h., wenn für $i \neq k$ jeder Vektor aus M_i orthogonal zu jedem Vektor aus M_k ist,

2. jedes Element $f \in H$ in der Form

$$f = h_1 + h_2 + \cdots + h_n + \cdots, \qquad h_n \in M_n,$$

dargestellt werden kann, wobei im Fall unendlich vieler Teilräume M_n die Reihe $\sum_n \|h_n\|^2$ konvergent sein soll. Wenn eine solche Darstellung des Elements f existiert, so ist sie eindeutig bestimmt, und es gilt

$$\|f\|^2 = \sum_n \|h_n\|^2,$$

wie man leicht nachprüfen kann.

Neben der direkten Summe von Teilräumen kann man auch eine direkte Summe endlich oder abzählbar vieler Hilberträume einführen. Und zwar wird die direkte Summe H zweier Hilberträume H_1 und H_2 folgendermaßen definiert. Als Elemente des Raumes H nimmt man alle Paare (h_1, h_2) mit $h_1 \in H_1$ und $h_2 \in H_2$, und das

Skalarprodukt zweier solcher Paare erklärt man durch

$$\big((h_1, h_2)\,(h_1', h_2')\big) = (h_1, h_1') + (h_2, h_2').$$

Die Paare der Gestalt $(h_1, 0)$ und die Paare der Gestalt $(0, h_2)$ bilden offenbar zwei zueinander orthogonale Teilräume in H; den ersten kann man auf natürliche Weise mit dem Raum H_1 und den zweiten mit dem Raum H_2 identifizieren.

Analog wird die Summe endlich vieler Räume definiert. Die Summe $H = \sum \oplus H_n$ von abzählbar vielen Räumen $H_1, H_2, \ldots, H_n, \ldots$ wird folgendermaßen erklärt: Die Elemente des Raumes H sind alle Folgen der Form

$$h = (h_1, h_2, \ldots, h_n, \ldots) \qquad (h_n \in H_n)$$

mit

$$\sum_n \|h_n\|^2 < \infty,$$

und als Skalarprodukt (h, g) von Elementen h und g aus H nimmt man die Reihe

$$\sum_n (h_n, g_n).$$

3.4.8. Charakteristische Eigenschaft unitärer Räume. Wir wollen folgende Frage untersuchen. R sei ein normierter Raum; welche zusätzlichen Bedingungen muß dann eine in R definierte Norm erfüllen, damit R zu einem unitären Raum wird, d. h., damit die Norm in R durch ein Skalarprodukt definiert werden kann. Oder, anders ausgedrückt, wie kann man unitäre Räume in der Klasse aller normierten Räume charakterisieren. Eine solche Charakterisierung liefert der folgende Satz.

Satz 8. *Ein normierter Raum R ist dann und nur dann ein unitärer Raum, wenn für je zwei Elemente f und g die Gleichung*

$$\|f + g\|^2 + \|f - g\|^2 = 2(\|f\|^2 + \|g\|^2) \tag{25}$$

erfüllt ist.

Beweis. Da $f + g$ und $f - g$ die Diagonalen eines Parallelogramms mit den Seiten f und g sind, drückt Gleichung (25) eine bekannte Eigenschaft des Parallelogramms im unitären Raum aus: *In einem Parallelogramm ist die Summe der Quadrate der Längen seiner Diagonalen gleich der Summe der Quadrate der Längen seiner Seiten.* Somit ist die Notwendigkeit der Bedingung klar.

Wir werden beweisen, daß sie auch hinreichend ist. Dazu setzen wir

$$(f, g) = \frac{1}{4}\,(\|f + g\|^2 - \|f - g\|^2) \tag{26}$$

und zeigen, daß die Funktion (26) allen Axiomen eines Skalarproduktes genügt, sofern (25) erfüllt ist. Da sich für $f = g$

$$(f, f) = \frac{1}{4}\,(\|2f\|^2 - \|f - f\|^2) = \|f\|^2 \tag{27}$$

ergibt, erzeugt dieses Skalarprodukt gerade die im Raum R gegebene Norm.

Zunächst ist aus (26) sofort ersichtlich, daß

$$(f, g) = \overline{(g, f)}$$

ist, d. h., Eigenschaft 1 eines Skalarproduktes ist erfüllt. Außerdem gilt wegen (27) auch Eigenschaft 4. Zum Nachweis von Eigenschaft 2 betrachten wir die Funktion dreier Vektoren

$$\Phi(f, g, h) = 4[(f + g, h) - (f, h) - (g, h)],$$

d. h.

$$\Phi(f, g, h) = \|f + g + h\|^2 - \|f + g - h\|^2 - \|f + h\|^2 + \|f - h\|^2 - \|g + h\|^2 + \|g - h\|^2, \tag{28}$$

und zeigen, daß diese identisch verschwindet. Wegen (25) gilt

$$\|f + g \pm h\|^2 = 2\|f \pm h\|^2 + 2\,\|g\|^2 - \|f \pm h - g\|^2.$$

Das setzen wir in (28) ein und erhalten

$$\Phi(f, g, h) = -\|f + h - g\|^2 + \|f - h - g\|^2 + \|f + h\|^2 - \|f - h\|^2 - \|g + h\|^2 + \|g - h\|^2. \tag{29}$$

Nun bilden wir das arithmetische Mittel von (28) und (29):

$$\Phi(f, g, h) = \frac{1}{2}\,(\|g + h + f\|^2 + \|g + h - f\|^2) - \frac{1}{2}\,(\|g - h + f\|^2 + \|g - h - f\|^2) - \|g + h\|^2 - \|g - h\|^2.$$

Nach (25) ist hier der erste Summand gleich

$$\|g + h\|^2 + \|f\|^2$$

und der zweite gleich

$$-\,\|g - h\|^2 - \|f\|^2.$$

Also gilt

$$\Phi(f, g, h) \equiv 0.$$

Wir weisen schließlich Eigenschaft 3, die Homogenität des Skalarprodukts nach. Dazu betrachten wir die Funktion

$$\varphi(c) = (cf, g) - c(f, g)$$

bei beliebigen festen f und g. Aus (26) folgt dann sofort

$$\varphi(0) = \frac{1}{4} (\|g\|^2 - \|g\|^2) = 0,$$

außerdem ist $\varphi(-1) = 0$ wegen $(-f, g) = -(f, g)$. Daher gilt für jede ganze Zahl n

$$\begin{aligned}(nf, g) &= (\operatorname{sgn} n(f + \cdots + f), g) \\ &= \operatorname{sgn} n[(f, g) + \cdots + (f, g)] \\ &= |n| \operatorname{sgn} n(f, g) = n(f, g),\end{aligned}$$

d. h. $\varphi(n) = 0$. Weiter ergibt sich für ganze Zahlen p und q mit $q \neq 0$

$$\left(\frac{p}{q} f, g\right) = p\left(\frac{1}{q} f, g\right) = \frac{p}{q} \cdot q\left(\frac{1}{q} f, g\right) = \frac{p}{q} (f, g),$$

d. h. $\varphi(c) = 0$ für alle rationalen Zahlen c. Da die Funktion φ stetig ist, folgt endlich

$$\varphi(c) \equiv 0.$$

Somit haben wir gezeigt, daß die Funktion (f, g) alle Eigenschaften eines Skalarproduktes besitzt.

Beispiele

1. Wir betrachten den n-dimensionalen Raum $\boldsymbol{R}_p{}^n$, in dem die Norm durch die Formel

$$\|x\|_p = \left(\sum_{k=1}^{n} |x_k|^p\right)^{1/p}$$

definiert ist. Für $p \geqq 1$ sind alle Axiome einer Norm erfüllt, jedoch ist $\boldsymbol{R}_p{}^n$ nur für $p = 2$ ein unitärer Raum. Wenn wir nämlich in $\boldsymbol{R}_p{}^n$ die beiden Vektoren

$$\begin{aligned}f &= (1, 1, 0, 0, \ldots, 0), \\ g &= (1, -1, 0, 0, \ldots, 0)\end{aligned}$$

betrachten, dann ist

$$\begin{aligned}f + g &= (2, 0, 0, \ldots, 0), \\ f - g &= (0, 2, 0, \ldots, 0)\end{aligned}$$

und daher

$$\|f\|_p = \|g\|_p = 2^{1/p}, \qquad \|f + g\|_p = \|f - g\|_p = 2,$$

so daß die Parallelogrammgleichung (25) für $p \neq 2$ nicht erfüllt ist.

2. Wir betrachten den Raum der stetigen Funktionen auf dem Intervall $[0, \pi/2]$. Für

$$f(t) = \cos t, \qquad g(t) = \sin t$$

gilt dann

$$\|f\| = \|g\| = 1$$

und

$$\|f + g\| = \max_{0 \leq t \leq \pi/2} |\cos t + \sin t| = \sqrt{2},$$

$$\|f - g\| = \max_{0 \leq t \leq \pi/2} |\cos t - \sin t| = 1.$$

Hieraus folgt

$$\|f + g\|^2 + \|f - g\|^2 \neq 2(\|f\|^2 + \|g\|^2).$$

Somit kann man die Norm im Raum $C[0, \pi/2]$ nicht mit Hilfe eines Skalarproduktes erklären. Es ist leicht zu sehen, daß auch der Raum $C[a, b]$ der stetigen Funktionen auf einem beliebigen Intervall $[a, b]$ kein unitärer Raum ist.

3.4.9. Komplexe unitäre Räume. Neben den reellen können auch komplexe unitäre Räume eingeführt werden (d. h. komplexe lineare Räume mit einem Skalarprodukt). Jedoch können die in 3.4.1. formulierten Axiome 1 bis 4 in einem komplexen Raum nicht gleichzeitig gelten. Aus 1 und 3 folgt nämlich

$$(\lambda x, \lambda x) = \lambda^2(x, x),$$

woraus wir für $\lambda = i$

$$(ix, ix) = -(x, x)$$

erhalten, d. h., die Skalarprodukte der Vektoren x und ix können nicht gleichzeitig positiv sein. Mit anderen Worten sind die Axiome 1 und 3 unvereinbar mit Axiom 4. Deshalb müssen die Axiome, mit deren Hilfe das Skalarprodukt definiert wurde, im komplexen Fall gegenüber dem reellen Fall etwas abgeändert werden. In einem komplexen Raum definieren wir das Skalarprodukt als (komplexwertige) Funktion zweier Vektoren, die folgenden Bedingungen genügt:

1. $(x, y) = \overline{(y, x)}$,
2. $(x_1 + x_2, y) = (x_1, y) + (x_2, y)$,
3. $(\lambda x, y) = \lambda(x, y)$,
4. $(x, x) \geqq 0$, wobei $(x, x) > 0$ ist für $x \neq 0$.

(Wir haben also im 1. Axiom eine Korrektur vorgenommen, während die drei anderen Axiome ohne Änderung beibehalten wurden.) Aus den Bedingungen 1 und 3 ergibt sich

$$(x, \lambda y) = \overline{(\lambda y, x)} = \overline{\lambda (y, x)} = \overline{\lambda}(x, y).$$

Ein wohlbekanntes Beispiel eines komplexen n-dimensionalen unitären Raumes ist der lineare Raum $\boldsymbol{C}^n$ (3.1.1., Beispiel 2), wenn das Skalarprodukt der Elemente

$$x = (x_1, x_2, \ldots, x_n) \quad \text{und} \quad y = (y_1, y_2, \ldots, y_n)$$

durch die Formel

$$(x, y) = \sum_{k=1}^{n} x_k \bar{y}_k$$

definiert wird. Bekanntlich ist jeder n-dimensionale komplexe euklidische Raum isomorph zu diesem Raum.

Als Beispiele unendlichdimensionaler komplexer unitärer Räume können dienen

1. der komplexe Raum l_2 aller Folgen

$$x = (x_1, x_2, \ldots, x_n, \ldots)$$

von komplexen Zahlen, für die

$$\sum_{n=1}^{\infty} |x_n|^2 < \infty$$

ist, mit dem Skalarprodukt

$$(x, y) = \sum_{n=1}^{\infty} x_n \bar{y}_n;$$

2. der Raum $C^2[a, b]$ aller stetigen komplexwertigen Funktionen auf dem Intervall $[a, b]$ mit dem Skalarprodukt

$$(f, g) = \int_a^b f(t)\,\overline{g(t)}\,dt.$$

In einem komplexen Raum wird die Länge (Norm) eines Vektors genau wie im reellen Fall durch die Formel

$$\|x\| = \sqrt{(x, x)}$$

definiert. Den Begriff des Winkels zwischen Vektoren führt man im komplexen Fall gewöhnlich nicht ein (da die Größe $\dfrac{(x, y)}{\|x\|\,\|y\|}$ im allgemeinen komplex ist und daher nicht Kosinus eines reellen Winkels sein kann). Jedoch der Begriff der Orthogonalität wird beibehalten: Zwei Elemente x und y heißen *orthogonal zueinander*, wenn $(x, y) = 0$ ist.

Ist $\{\varphi_n\}$ ein orthogonales System im komplexen unitären Raum R und f ein beliebiges Element aus R, so werden die Zahlen

$$a_n = \frac{1}{\|\varphi_n\|^2} (f, \varphi_n)$$

wie im reellen Fall *Fourierkoeffizienten* genannt, und die Reihe

$$\sum_n a_n \varphi_n$$

heißt *Fourierreihe* des Elements f bezüglich des Systems $\{\varphi_n\}$. Dabei gilt die Besselsche Ungleichung

$$\sum_n \|\varphi_n\|^2 \, |a_n|^2 \leqq (f, f).$$

Ist insbesondere das System $\{\varphi_n\}$ orthonormiert, so werden die Fourierkoeffizienten bezüglich dieses Systems durch die Formel

$$c_n = (f, \varphi_n)$$

gegeben, und die Besselsche Ungleichung nimmt die Gestalt

$$\sum_n |c_n|^2 \leqq (f, f)$$

an. Ein vollständiger separabler komplexer unitärer Raum unendlicher Dimension heißt *komplexer Hilbertraum*. Auch für komplexe Hilberträume gilt der Isomorphiesatz:

Satz 9. *Alle komplexen Hilberträume sind zueinander isomorph.*

Die einfachste Realisierung eines solchen Raumes ist der komplexe Raum l_2. Eine andere Realisierung des komplexen Hilbertraumes durch einen Funktionenraum werden wir in Kapitel 6 kennenlernen.

Der Leser möge nachprüfen, daß alle Sätze, die für reelle unitäre Räume und insbesondere Hilberträume gezeigt worden sind, (mit unbedeutenden Abänderungen wegen der abweichenden Definition des Skalarproduktes) auch für komplexe Räume gelten.

3.5. Topologische lineare Räume

3.5.1. Definition und Beispiele. Die Angabe einer Norm ist nur eine der möglichen Arten, in einem linearen Raum eine Topologie einzuführen. Die Entwicklung der Theorie der Distributionen (von denen im nächsten Kapitel die Rede sein wird) und auch anderer Teilgebiete der Funktionalanalysis zeigte, daß in vielen Fällen die Betrachtung linearer Räume nützlich ist, in denen die Topologie nicht mit Hilfe einer Norm, sondern auf irgendeine andere Weise eingeführt wurde.

Definition 1. Eine Menge E heißt *topologischer linearer* (oder auch *linearer topologischer*) Raum, wenn folgendes gilt:

I. E ist ein (reeller oder komplexer) linearer Raum.

II. E ist ein topologischer Raum.

III. Die Addition und die Skalarmultiplikation in E sind bezüglich der in E gegebenen Topologie stetig. Genauer besagt die letzte Bedingung:

1. Wenn $z_0 = x_0 + y_0$ ist, dann gibt es zu jeder Umgebung U des Punktes z_0 Umgebungen V und W von x_0 bzw. y_0, so daß $x + y$ für $x \in V$, $y \in W$ zu U gehört.

2. Wenn $\alpha_0 x_0 = y_0$ ist, existiert zu jeder Umgebung U des Punktes y_0 eine Umgebung V des Punktes x_0 und eine Zahl $\varepsilon > 0$, so daß αx für $|\alpha - \alpha_0| < \varepsilon$ und $x \in V$ zu U gehört.

Aus diesem Zusammenhang zwischen den algebraischen Operationen und der Topologie in einem topologischen linearen Raum folgt, daß die Topologie in einen solchen Raum bereits vollständig durch das *System aller Nullumgebungen* bestimmt wird. Ist nämlich x ein Punkt des linearen topologischen Raumes E und U eine Umgebung des Nullpunktes in E, so ist die in den Punkt x verschobene Umgebung $U + x$ eine Umgebung des Punktes x. Offenbar kann jede Umgebung eines beliebigen Punktes $x \in E$ auch auf diese Weise erhalten werden.

Aus der Stetigkeit der Addition und der Skalarmultiplikation in einem topologischen linearen Raum ergeben sich unmittelbar die folgenden Aussagen:

1. *Sind U, V offene Mengen in E, dann ist auch die Menge $U + V$* (d. h. die Gesamtheit aller Elemente der Form $x + y$, $x \in U$, $y \in V$) *offen.*

2. *Ist U offen, dann ist auch die Menge λU* (d. h. die Gesamtheit aller Elemente der Form λx, $x \in U$) *für jedes $\lambda \neq 0$ offen.*

3. *Ist F eine abgeschlossene Menge in E, so ist auch λF für jedes λ abgeschlossen.*

Beispiele

1. Zu den topologischen linearen Räumen gehören zunächst einmal alle normierten Räume, denn aus den Eigenschaften der Norm folgt sofort, daß die Addition und die Skalarmultiplikation in der durch die Norm erzeugten Topologie stetig sind.

2. Wir definieren im Raum $\boldsymbol{R}^\infty$ aller Zahlenfolgen $x = (x_1, x_2, \ldots, x_n, \ldots)$ folgendermaßen ein System von Nullumgebungen. Jeweils endlich viele natürliche Zahlen $k_1, \ldots, k_r$ und eine positive Zahl $\varepsilon > 0$ bestimmen eine Umgebung $U(k_1, \ldots, k_r, \varepsilon)$, die aus allen Punkten $x \in \boldsymbol{R}^\infty$ mit

$$|x_{k_i}| < \varepsilon, \qquad i = 1, 2, \ldots, r,$$

besteht. Man prüft leicht nach, daß die Angabe dieses Systems von Nullumgebungen den Raum $\boldsymbol{R}^\infty$ zu einem topologischen linearen Raum macht. (Neben $\boldsymbol{R}^\infty$ kann man auch den Raum $\boldsymbol{C}^\infty$ aller komplexen Folgen betrachten.)

3. Es sei $K[a, b]$ der Raum der unendlich oft differenzierbaren Funktionen[1]) auf dem Intervall $[a, b]$. In $K[a, b]$ definieren wir mit Hilfe des folgenden Systems von

[1]) Das sind Funktionen mit Ableitungen beliebig hoher Ordnung.

Nullumgebungen eine Topologie. Jede natürliche Zahl m und jede positive Zahl $\varepsilon > 0$ bestimmen eine Nullumgebung $U_{m,\varepsilon}$, die aus allen Funktionen φ mit

$$|\varphi^{(k)}(x)| < \varepsilon, \qquad k = 0, 1, 2, \ldots, m,$$

besteht; dabei ist $\varphi^{(k)}$ die k-te Ableitung der Funktion φ.

Die Forderung, daß die Topologie in einem topologischen linearen Raum mit den darin erklärten linearen Operationen verträglich sein soll, legt der Topologie hinreichend starke Einschränkungen auf. *In einem topologischen linearen Raum E besitzen ein Punkt x und eine diesen Punkt nicht enthaltende abgeschlossene Menge stets disjunkte Umgebungen.*

Zum Beweis dieser Behauptung genügt es, den Punkt $x = 0$ und eine beliebige ihn nicht enthaltende abgeschlossene Menge F zu betrachten. Wenn wir $U = E \setminus F$ setzen, gibt es wegen der Stetigkeit der Subtraktion eine Nullumgebung W mit $W - W \subset U$. Als eine solche Umgebung W kann man etwa den Durchschnitt zweier Nullumgebungen W_1 und W_2 nehmen, für die $x - y$ zu U gehört, falls $x \in W_1$ und $y \in W_2$ ist. Wir wollen nachweisen, daß der Abschluß der Umgebung W in U enthalten ist. Dazu sei $y \in [W]$. Dann enthält jede Umgebung des Punktes y, insbesondere also $y + W$, einen Punkt z aus W. Folglich gilt $y + z \in W$, d. h. also $y \in W - W \subset U$, was gerade behauptet worden war. Die Mengen W und $E \setminus [W]$ sind dann die gesuchten Umgebungen des Punktes 0 bzw. der Menge F.

Ein topologischer Raum heißt *separiert*, wenn er dem ersten Trennungsaxiom genügt, d. h., wenn alle einpunktigen Mengen abgeschlossen sind. Offenbar ist ein linearer topologischer Raum dann und nur dann separiert, wenn der Durchschnitt aller Nullumgebungen keine von Null verschiedenen Elemente enthält. Topologische Räume, die dem ersten und dem dritten Trennungsaxiom genügen, haben wir in Kapitel 2 *regulär* genannt. Nach dem im vorhergehenden Absatz Gezeigten folgt dann, daß *ein separierter linearer topologischer Raum regulär ist.*

In normierten Räumen spielt der Begriff der beschränkten Menge eine wichtige Rolle. Obwohl dieser Begriff dort mit Hilfe der Norm eingeführt wird, kann man ihn auch in natürlicher Weise für beliebige topologische lineare Räume formulieren.

Eine Menge M im topologischen Raum E nennen wir *beschränkt*, wenn für jede Nullumgebung U ein $n > 0$ existiert, so daß $\lambda U \supset M$ ist für $|\lambda| \geqq n$.

Es ist klar, daß dieser Beschränktheitsbegriff für normierte Räume mit der Beschränktheit bezüglich der Norm übereinstimmt (d. h. mit der Möglichkeit, eine gegebene Menge im Innern einer Kugel $\|x\| \leqq R$ unterzubringen). Ein Raum E heißt *lokalbeschränkt*, wenn in E wenigstens eine nichtleere offene beschränkte Menge existiert. Jeder normierte Raum ist lokalbeschränkt. Der in Beispiel 2 angegebene Raum $\boldsymbol{R}^\infty$ ist ein Beispiel eines nicht lokalbeschränkten Raumes (der Leser möge das beweisen).

Aufgaben

1. E sei ein topologischer linearer Raum, dann sind die folgenden Behauptungen zu beweisen.

a) Eine Menge $M \subset E$ ist dann und nur dann beschränkt, wenn für jede Folge $\{x_n\} \subset M$ und jede Nullfolge $\{\varepsilon_n\}$ von positiven Zahlen die Folge $\varepsilon_n x_n$ gegen Null strebt.

b) Jede konvergente Folge $\{x_n\} \subset E$ ist eine beschränkte Menge.

c) Wenn E lokalbeschränkt ist, erfüllt E das erste Abzählbarkeitsaxiom.

Gilt das erste Abzählbarkeitsaxiom im Raum $\boldsymbol{R}^\infty$?

2. Wir sagen, daß eine Menge M in einem topologischen linearen Raum von der Nullumgebung U *absorbiert* wird, wenn eine Zahl n mit $nU \supset M$ existiert. Man beweise, daß in einem lokalbeschränkten Raum ein Fundamentalsystem von Nullumgebungen existiert, die einander absorbieren, und gebe ein solches System in einem normierten Raum an.

3.5.2. Lokalkonvexe Räume. Ein beliebiger topologischer linearer Raum kann Eigenschaften besitzen, die sehr stark von den gewohnten Eigenschaften unitärer oder normierter Räume abweichen. Die sogenannten *lokalkonvexen* Räume bilden eine wichtige Klasse von Räumen, die allgemeiner sind als die normierten Räume, für die aber viele Eigenschaften dieser Räume erhalten bleiben.

Definition 2. Ein topologischer linearer Raum heißt *lokalkonvex,* wenn darin jede nichtleere offene Menge eine nichtleere konvexe offene Teilmenge enthält.

Wenn der Raum E lokalkonvex ist, gibt es zu jedem Punkt $x \in E$ und jeder Umgebung U von x eine konvexe Umgebung V von x mit $V \subset U$. Es reicht aus, die Richtigkeit dieser Behauptung für den Punkt $x = 0$ nachzuweisen. Dazu sei U irgendeine Nullumgebung. Dann gibt es eine Nullumgebung V mit $V - V \subset U$. Da E lokalkonvex ist, enthält V eine nichtleere konvexe offene Teilmenge $V' \subset V$. Es sei $y \in V'$, dann ist $V' - y$ eine konvexe Nullumgebung, die in U enthalten ist.

Jeder normierte Raum ist lokalkonvex. Denn jede nichtleere offene Teilmenge in einem normierten Raum enthält eine Kugel. *Somit ist jeder normierte Raum lokalbeschränkt und lokalkonvex.* Man kann zeigen, daß die Klasse der Räume mit diesen beiden Eigenschaften im wesentlichen auch nur die normierten Räume enthält.

Ein topologischer linearer Raum E heißt *normierbar,* wenn die Topologie in E mit Hilfe einer Norm erklärt werden kann. Dann gilt folgender Satz: *Jeder separierte lokalkonvexe und lokalbeschränkte topologische lineare Raum ist normierbar.*

Aufgaben

1. Man zeige, daß eine offene Menge U in einem topologischen linearen Raum genau dann konvex ist, wenn $U + U = 2U$ ist.

2. E sei ein linearer Raum. Eine Menge $U \subset E$ heißt *symmetrisch,* wenn mit $x \in U$ auch $-x \in U$ gilt. B sei das System aller konvexen symmetrischen Teilmengen des Raumes E, die mit ihrem offenen Kern (vgl. 3.2.1.) übereinstimmen. Dann ist die Gültigkeit folgender Behauptungen zu zeigen:

a) Das System B ist ein Erzeugendensystem von Nullumgebungen für eine lokalkonvexe separierte Topologie E (diese Topologie heißt *offen-konvex*).

b) Die offen-konvexe Topologie ist die gröbste aller lokalkonvexen Topologien, in der die linearen Operationen in E stetig sind.

c) Jedes lineare Funktional auf E ist bezüglich der offen-konvexen Topologie stetig.

3.5.3. Abzählbar-normierte Räume. Eine für die Analysis sehr wichtige Klasse von topologischen linearen Räumen bilden die sogenannten *abzählbar-normierten Räume*. Um die entsprechende Definition formulieren zu können, benötigen wir erst noch einen weiteren Begriff.

In einem linearen Raum E seien zwei Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ gegeben. Diese Normen heißen *koordiniert*, wenn jede Folge aus E, die in der einen Norm gegen den Grenzwert $x \in E$ konvergiert und bezüglich der anderen Norm eine Fundamentalfolge bildet, auch in der zweiten Norm gegen denselben Grenzwert x konvergiert.

Die Norm $\|\cdot\|_1$ nennt man *stärker* als die Norm $\|\cdot\|_2$, wenn eine Konstante $c > 0$ existiert, so daß $\|x\|_1 \geqq c\,\|x\|_2$ für alle $x \in E$ ist.

Wenn die erste Norm stärker als die zweite und die zweite stärker als die erste ist, heißen die beiden Normen *äquivalent*. Zwei Normen heißen *vergleichbar*, wenn eine von ihnen stärker als die andere ist.

Definition 3. Ein linearer Raum E heißt *abzählbar-normiert*, wenn in E ein abzählbares System paarweise koordinierter Normen $\|\cdot\|_n$ gegeben ist. Jeder abzählbar-normierte Raum wird zu einem topologischen linearen Raum, wenn man als erzeugendes Nullumgebungssystem die Gesamtheit aller Mengen

$$U_{r,\varepsilon} = \{x\colon x \in E, \|x\|_1 < \varepsilon, \ldots, \|x\|_r < \varepsilon\}$$

(r natürliche Zahl, $\varepsilon > 0$) nimmt.

Der Leser möge als Übung nachweisen, daß dieses System von Nullumgebungen tatsächlich in E eine Topologie definiert, in der die Addition und die Skalarmultiplikation stetig sind.

Jeder abzählbar-normierte Raum genügt dem ersten Abzählbarkeitsaxiom, da man das System aller Nullumgebungen $U_{r,\varepsilon}$ durch das abzählbare Teilsystem aller Nullumgebungen $U_{r,1/n}$ ersetzen kann (ohne dabei die Topologie zu verändern). Darüber hinaus kann man die Topologie in einem abzählbar-normierten Raum auch mit Hilfe einer Metrik definieren, z. B. durch

$$\varrho(x, y) = \sum_{n=1}^{\infty} 2^{-n} \frac{\|x-y\|_n}{1 + \|x-y\|_n}, \qquad x, y \in E. \tag{1}$$

Wir überlassen dem Leser den Nachweis, daß die Funktion $\varrho(x, y)$ alle Abstandsaxiome erfüllt und translationsinvariant ist (d. h. $\varrho(x+z, y+z) = \varrho(x, y)$, für $x, y, z \in E$) und daß die davon erzeugte Topologie mit der ursprünglichen übereinstimmt. Somit können wir von der Vollständigkeit eines abzählbar-normierten Raumes sprechen, wenn wir darunter die Vollständigkeit bezüglich der oben eingeführten Metrik verstehen. Wir bemerken noch, daß eine Folge $\{x_k\}$ dann und nur dann Fundamentalfolge bezüglich der Metrik (1) ist, wenn sie Fundamentalfolge in jeder der Normen $\|\cdot\|_n$ ist, und daß sie genau dann in dieser Metrik gegen ein

Element $x \in E$ konvergiert, wenn sie in jeder Norm $\|\cdot\|_n$ gegen x konvergiert. Mit anderen Worten bedeutet die Vollständigkeit eines abzählbar-normierten Raumes, daß in ihm jede Folge konvergiert, die Fundamentalfolge bezüglich aller Normen $\|\cdot\|_n$ ist.

Beispiele

1. Ein wichtiges Beispiel eines abzählbar-normierten Raumes ist der oben betrachtete Raum $K[a, b]$ aller auf dem Intervall $[a, b]$ unendlich oft differenzierbaren Funktionen, wenn man die Norm $\|\cdot\|_m$ in diesem Raum durch die Formel

$$\|f\|_m = \sup_{\substack{a \leqq t \leqq b \\ 0 \leqq k \leqq m}} |f^{(k)}(t)|$$

definiert. Offenbar sind alle diese Normen untereinander koordiniert und definieren in $K[a, b]$ genau die oben beschriebene Topologie.

2. Es sei S der Raum aller der unendlich oft differenzierbaren Funktionen auf der Zahlengeraden, die zusammen mit allen ihren Ableitungen im Unendlichen schneller als jede Potenz von $1/|t|$ gegen Null streben (d. h. die der Bedingung $t^k f^{(q)}(t) \to 0$ für $|t| \to \infty$ bei beliebig fixiertem k und q genügen). In diesem Raum definieren wir ein abzählbares System von Normen, indem wir

$$\|f\|_m = \sup_{\substack{0 \leqq k, q \leqq m \\ -\infty < t < \infty}} |t^k f^{(q)}(t)|, \qquad m = 0, 1, 2, \ldots,$$

setzen. Es läßt sich leicht nachprüfen, daß diese Normen untereinander koordiniert sind. Somit ist S ein abzählbar-normierter Raum.

3. Einen wichtigen Spezialfall abzählbar-normierter Räume bilden die sogenannten abzählbar-Hilbertschen Räume. Es sei H ein linearer Raum, in dem ein abzählbares System $(\varphi, \psi)_n$ von Skalarprodukten gegeben ist. Dabei setzen wir voraus, daß die entsprechenden Normen $\|\varphi\|_n = \sqrt{(\varphi, \varphi)_n}$ untereinander koordiniert sind. Wenn ein solcher Raum vollständig ist, heißt er *abzählbar-Hilbertscher Raum.*

4. Als konkretes Beispiel eines abzählbar-Hilbertschen Raumes kann der folgende Raum dienen. Es sei Φ die Gesamtheit aller Zahlenfolgen $\{x_n\}$, für die die Reihe

$$\sum_{n=1}^{\infty} n^k x_n^2$$

bei beliebigem $k \geqq 0$ konvergiert. Wenn wir

$$\|x\|_k = \sqrt{\sum_{n=1}^{\infty} n^k x_n^2}$$

setzen, erhalten wir in diesem Raum ein abzählbares System von Normen. Unschwer überzeugt man sich davon, daß diese Normen untereinander koordiniert sind, und daß

Φ im oben angegebenen Sinne vollständig ist. Offensichtlich kann man jede der Normen mit Hilfe des Skalarprodukts

$$(x, y)_k = \sum_{n=1}^{\infty} n^k x_n y_n$$

definieren, d. h., Φ ist ein abzählbar-Hilbertscher Raum. Er wird *Raum der schnell fallenden Folgen* genannt.

Wenn E ein abzählbar-normierter Raum ist, kann man annehmen, daß die darin erklärten Normen $\|\cdot\|_k$ der Bedingung

$$\|x\|_k \leqq \|x\|_l \quad \text{für} \quad k < l \tag{2}$$

genügen. Denn sonst könnte man die Normen $\|x\|_k$ durch die Normen

$$\|x\|_k' = \sup (\|x\|_1, \|x\|_2, \ldots, \|x\|_k)$$

ersetzen, die in E dieselbe Topologie bestimmen wie das System der ursprünglichen Normen. Durch Vervollständigung des Raumes E in jeder der Normen $\|\cdot\|_k$ erhalten wir ein System vollständiger normierter Räume E_k. Dabei folgt aus der Beziehung (2) und der Koordiniertheit der Normen, daß die natürliche Einbettung

$$E_k \supset E_l \quad \text{für} \quad k < l$$

besteht. Somit kann man jedem abzählbar-normierten Raum E eine fallende Kette vollständiger normierter Räume

$$E_1 \supset E_2 \supset \cdots \supset E_k \supset \ldots, \qquad \bigcap_{k=1}^{\infty} E_k \supset E,$$

zuordnen. Man kann zeigen, daß der Raum E genau dann vollständig ist, wenn $E = \bigcap_{k=1}^{\infty} E_k$ ist (Beweis!). So ist zum Beispiel der Raum $K[a, b]$ der unendlich oft differenzierbaren Funktionen auf dem Intervall $[a, b]$ der Durchschnitt der vollständigen normierten Räume D^n $(n = 0, 1, 2, \ldots)$, dabei ist D^n der Raum aller n-mal stetig differenzierbaren Funktionen mit der Norm

$$\|f\|_n = \sup_{\substack{a \leqq t \leqq b \\ 0 \leqq k \leqq n}} |f^{(k)}(t)|.$$

Die Theorie der linearen normierten Räume ist in den dreißiger Jahren, hauptsächlich in den Arbeiten von Banach, entwickelt worden. Damals war man der Ansicht, daß diese Klasse von Räumen umfangreich genug ist, um alle konkreten Bedürfnisse der Analysis zu erfüllen. In der Folgezeit wurde jedoch klar, daß das nicht so ist. Es zeigte sich, daß für eine Reihe von Problemen solche Räume wie der Raum der unendlich oft differenzierbaren Funktionen, der Raum $\boldsymbol{R}^\infty$ aller Zahlenfolgen und andere Räume wichtig sind, in denen die jeweils natürliche Topologie nicht mit Hilfe irgendeiner Norm erklärt werden kann. Somit sind topologische, aber nicht normierte lineare Räume durchaus keine „pathologischen" Fälle. Vielmehr stellen einige dieser Räume genauso natürliche und wichtige Verallgemeinerungen des endlichdimensionalen euklidischen Raumes dar wie beispielsweise der Hilbertraum.

4. Lineare Funktionale und lineare Operatoren

4.1. Stetige lineare Funktionale

4.1.1. Stetige lineare Funktionale in topologischen linearen Räumen. In 3.1. haben wir bereits Funktionale auf einem linearen Raum betrachtet. Unter den Funktionalen auf einem topologischen linearen Raum sind die stetigen Funktionale von grundlegendem Interesse. Wie üblich heißt ein Funktional f auf einem Raum E *stetig*, wenn zu jedem $\varepsilon > 0$ und jedem $x_0 \in E$ eine Umgebung U des Elementes x existiert, so daß

$$|f(x) - f(x_0)| < \varepsilon \quad \text{für} \quad x \in U \tag{1}$$

ist. Diese Definition bezieht sich insbesondere auch auf lineare Funktionale.

Ist E ein endlichdimensionaler topologischer linearer Raum, so ist jedes lineare Funktional auf E automatisch stetig. Im allgemeinen folgt jedoch aus der Linearität eines Funktionals nicht dessen Stetigkeit.

Wenn ein lineares Funktional f in einem Punkt $x \in E$ stetig ist, so ist f überall auf E stetig.

Um das zu zeigen, nehmen wir an, daß y ein beliebiger Punkt in E und $\varepsilon > 0$ ist. Wir wählen eine Umgebung U des Punktes x, so daß die Bedingung (1) erfüllt ist. Dann ist die verschobene Umgebung

$$V = U + (y - x)$$

die gesuchte Umgebung des Punktes y, denn für jedes $z \in V$ gilt $z + x - y \in U$ und folglich

$$|f(z) - f(y)| = |f(z - y + x) - f(x)| < \varepsilon.$$

Somit reicht es aus, die Stetigkeit eines linearen Funktionals in einem Punkt nachzuweisen, zum Beispiel im Punkt 0.

Wenn E ein Raum ist, der dem ersten Abzählbarkeitsaxiom genügt, kann man die Stetigkeit eines linearen Funktionals auf E mit Hilfe von Folgen formulieren: Ein Funktional f heißt stetig im Punkt $x \in E$, wenn aus $x_n \to x$ auch $f(x_n) \to f(x)$ folgt. Der Nachweis der Gleichwertigkeit dieser Stetigkeitsdefinition mit der oben angegebenen (im Fall der Gültigkeit des ersten Abzählbarkeitsaxioms) wird dem Leser überlassen.

Satz 1. *Für die Stetigkeit eines linearen Funktionals f auf E ist notwendig und hinreichend, daß eine Nullumgebung in E existiert, auf der f beschränkt ist.*

Beweis. Wenn das Funktional f im Punkt 0 stetig ist, existiert zu jedem $\varepsilon > 0$ eine Nullumgebung, auf der

$$|f(x)| < \varepsilon$$

ist. Umgekehrt sei U eine Nullumgebung mit

$$|f(x)| \leqq C \quad \text{für} \quad x \in U,$$

und es sei $\varepsilon > 0$. Dann ist $\frac{\varepsilon}{C} U$ eine Nullumgebung, auf der $|f(x)| < \varepsilon$ ist.

Aufgabe. E sei ein topologischer linearer Raum; man zeige die Gültigkeit der folgenden Behauptungen:

a) Ein lineares Funktional auf E ist dann und nur dann stetig, wenn eine offene Menge $U \subset E$ und eine Zahl t mit $t \notin f(U)$ existieren; dabei ist $f(U)$ die Menge aller Werte von $f(x)$, $x \in U$.

b) Ein lineares Funktional ist genau dann stetig, wenn sein Nullraum $\{x: f(x) = 0\}$ in E abgeschlossen ist.

c) Wenn jedes lineare Funktional auf E stetig ist, stimmt die Topologie in E mit der offenkonvexen Topologie (vgl. Aufgabe 2 auf S. 172) überein.

d) Wenn E unendlichdimensional und normierbar ist, dann existiert auf E ein nicht-stetiges lineares Funktional (man nutze die Existenz einer Hamelbasis in E aus; vgl. die Aufgabe auf S. 127).

e) In E existiere ein erzeugendes System von Nullumgebungen, dessen Mächtigkeit die algebraische Dimension des Raumes E (d. h. die Mächtigkeit einer Hamelbasis in E; vgl. die Aufgabe auf S. 127) nicht übertrifft. Dann existiert in E ein nicht-stetiges lineares Funktional.

f) Notwendig für die Stetigkeit eines linearen Funktionals f auf E ist die Beschränktheit von f auf jeder beschränkten Menge. Wenn E dem ersten Abzählbarkeitsaxiom genügt, ist diese Bedingung auch hinreichend.

4.1.2. Lineare Funktionale auf normierten Räumen. Der betrachtete Raum E sei jetzt normiert. Nach Satz 1 ist jedes stetige lineare Funktional f auf einer gewissen Nullumgebung beschränkt. Da in normierten Räumen jede Nullumgebung eine Kugel enthält, ist f also auf einer gewissen Kugel beschränkt. Wegen der Linearität des Funktionals ist das gleichwertig mit der Beschränktheit auf jeder Kugel, insbesondere auf der Einheitskugel $\|x\| \leqq 1$. Umgekehrt folgt aus der Beschränktheit des Funktionals f auf der Einheitskugel wiederum nach Satz 1 seine Beschränktheit (da das Innere dieser Kugel eine offene Menge ist).

Somit ist ein lineares Funktional in einem normierten Raum dann und nur dann stetig, wenn seine Werte auf der Einheitskugel beschränkt sind.

Es sei f ein stetiges lineares Funktional im normierten Raum E. Die Zahl

$$\|f\| = \sup_{\|x\| \leqq 1} |f(x)|, \tag{2}$$

d. h. die kleinste obere Schranke der Funktionswerte $|f(x)|$ auf der Einheitskugel des Raumes E, nennen wir die *Norm* des Funktionals f. Wir vermerken die folgenden fast trivialen Eigenschaften von $\|f\|$:

1. $$\|f\| = \sup_{x \neq 0} \frac{|f(x)|}{\|x\|};$$

das ergibt sich sofort aus der für $x \neq 0$ gültigen Beziehung

$$\frac{|f(x)|}{\|x\|} = \left|f\left(\frac{x}{\|x\|}\right)\right|.$$

2. Für jedes $x \in E$ gilt

$$|f(x)| \leqq \|f\| \cdot \|x\|. \tag{3}$$

Denn für $x \neq 0$ gehört das Element $\frac{x}{\|x\|}$ zur Einheitskugel, und daher ist nach Definition der Norm eines Funktionals

$$\left|f\left(\frac{x}{\|x\|}\right)\right| = \frac{|f(x)|}{\|x\|} \leqq \|f\|,$$

woraus (3) folgt. Für $x = 0$ steht rechts und links in (3) Null.

Aufgabe. Es sei $C \geqq 0$ eine Zahl mit

$$|f(x)| \leqq C\,\|x\| \tag{4}$$

für jedes x. Man weise die Beziehung $\|f\| = \inf C$ nach, wobei das Infimum über alle C genommen wird, die die Ungleichung (4) erfüllen.

Wir betrachten jetzt Beispiele linearer Funktionale in normierten Räumen.

1. Es sei a ein fester Vektor im reellen n-dimensionalen euklidischen Raum $\boldsymbol{R}^n$. Dann stellt das Skalarprodukt

$$f(x) = (x, a), \qquad x \in \boldsymbol{R}^n,$$

offenbar ein lineares Funktional auf dem $\boldsymbol{R}^n$ dar. Auf Grund der Schwarzschen Ungleichung gilt

$$|f(x)| = |(x, a)| \leqq \|x\| \cdot \|a\|, \tag{5}$$

folglich ist dieses Funktional beschränkt und somit auch stetig auf dem $\boldsymbol{R}^n$. Aus Ungleichung (5) erhalten wir

$$\frac{|f(x)|}{\|x\|} \leqq \|a\|.$$

Da die rechte Seite dieser Ungleichung nicht von x abhängt, folgt daraus

$$\sup \frac{|f(x)|}{\|x\|} \leqq \|a\|,$$

d. h. $\|f\| \leqq \|a\|$. Für $x = a$ erhalten wir

$$|f(a)| = (a, a) = \|a\|^2, \quad \text{d. h.} \quad \frac{|f(a)|}{\|a\|} = \|a\|.$$

Daher gilt sogar

$$\|f\| = \|a\|.$$

2. Das bestimmte Integral

$$I(x) = \int_a^b x(t)\, dt$$

einer auf $[a, b]$ stetigen Funktion $x(t)$ stellt ein lineares Funktional im Raum $C[a, b]$ dar. Dieses Funktional ist beschränkt, und seine Norm ist $b - a$. Denn es gilt

$$|I(x)| = \left| \int_a^b x(t)\, dt \right| \leqq \max |x(t)|\, (b - a) = \|x\|\, (b - a),$$

wobei für $x \equiv \text{const}$ die Gleichheit erreicht wird.

3. Wir betrachten ein allgemeineres Beispiel. Es sei $y_0(t)$ eine feste auf $[a, b]$ stetige Funktion. Dann setzen wir für jede Funktion $x(t) \in C[a, b]$

$$F(x) = \int_a^b x(t)\, y_0(t)\, dt.$$

Dieses Funktional ist linear und wegen

$$|F(x)| = \left| \int_a^b x(t)\, y_0(t)\, dt \right| \leqq \|x\| \int_a^b |y_0(t)|\, dt \tag{6}$$

auch beschränkt. Auf Grund der Linearität und Beschränktheit ist F also stetig. Aus (6) ergibt sich eine Abschätzung der Norm von F:

$$\|F\| \leqq \int_a^b |y_0(t)|\, dt.$$

(Man zeige, daß hier sogar die Gleichheit gilt!)

4. Wir betrachten im Raum $C[a, b]$ das schon in 3.1.5. erwähnte lineare Funktional

$$\delta_{t_0}(x) = x(t_0).$$

Für eine Funktion $x(t)$ ist der Wert dieses Funktionals gleich dem Funktionswert von $x(t)$ im Punkt t_0. Offenbar ist

$$|x(t_0)| \leqq \|x\|,$$

wobei für $x \equiv \text{const}$ die Gleichheit gilt. Hieraus ergibt sich sofort, daß die Norm des Funktionals δ_{t_0} gleich 1 ist.

5. In einem beliebigen unitären Raum X kann man ebenso wie im $\boldsymbol{R}^n$ ein lineares Funktional definieren, indem man für jedes $x \in X$

$$F(x) = (x, a)$$

mit einem festen Element $a \in X$ setzt. Wie im $\boldsymbol{R}^n$ ist die Beziehung

$$\|F\| = \|a\|$$

auch hier leicht nachzuprüfen.

Im weiteren werden wir nur stetige lineare Funktionale betrachten und das Wort „stetig" der Kürze halber weglassen.

Dem Begriff der Norm eines linearen Funktionals kann man folgende anschauliche Interpretation geben. Wir haben schon gesehen (3.1.), daß man jedem linearen Funktional eine Hyperebene L zuordnen kann, die durch die Gleichung

$$f(x) = 1$$

bestimmt wird. Wir wollen den Abstand d dieser Hyperebene vom Punkt 0 berechnen. Nach Definition ist $d = \inf\limits_{f(x)=1} \|x\|$. Aus der Ungleichung

$$|f(x)| \leqq \|f\| \cdot \|x\|$$

erhalten wir auf der Hyperebene $f(x) = 1$ die Abschätzung $\|x\| \geqq \frac{1}{\|f\|}$, daraus folgt $d \geqq \frac{1}{\|f\|}$. Andererseits kann man nach Definition der Norm von f zu jedem $\varepsilon > 0$ ein Element x_ε mit $f(x_\varepsilon) = 1$ und

$$1 > (\|f\| - \varepsilon) \|x_\varepsilon\|$$

finden; daher ist

$$d = \inf_{f(x)=1} \|x\| < \frac{1}{\|f\| - \varepsilon}.$$

Da $\varepsilon > 0$ beliebig war, erhalten wir

$$d = \frac{1}{\|f\|},$$

d. h., *die Norm eines linearen Funktionals ist der reziproke Wert des Abstands der Hyperebene $f(x) = 1$ vom Punkt* 0.

4.1.3. Der Satz von Hahn-Banach in normierten Räumen. In 3.2. haben wir den Satz von HAHN-BANACH allgemein bewiesen. Danach kann jedes lineare Funktional, das auf einem Teilraum L des linearen Raumes E definiert ist und dort der Bedingung

$$|f_0(x)| \leqq p(x) \tag{7}$$

genügt (p war dabei ein festes konvexes Funktional auf E), unter Beibehaltung dieser Bedingung auf ganz E fortgesetzt werden. Für normierte Räume kann man diesen Satz wie folgt formulieren.

Es sei E ein reeller normierter Raum, L ein Teilraum von E und f_0 ein beschränktes lineares Funktional auf L. Dann kann dieses lineare Funktional zu einem beschränkten linearen Funktional f auf dem ganzen Raum E fortgesetzt werden unter Beibehaltung seiner Norm, d. h.

$$\|f_0\|_{\text{auf}\, L} = \|f\|_{\text{auf}\, E}.$$

Wenn nämlich

$$\|f_0\|_{\text{auf}\, L} = k$$

ist, wird durch $p(x) = k\|x\|$ offenbar ein konvexes Funktional auf E mit

$$\|f_0(x)\| \leqq p(x)$$

für $x \in L$ definiert, und durch Anwendung des (allgemeinen) Satzes von HAHN-BANACH erhalten wir das gewünschte Resultat.

Diese Form des Satzes von HAHN-BANACH gestattet folgende geometrische Interpretation. Die Gleichung

$$f_0(x) = 1 \tag{8}$$

definiert im Teilraum L eine Hyperebene, deren Abstand vom Nullpunkt $\dfrac{1}{\|f_0\|}$ beträgt. Wenn wir das Funktional f_0 unter Beibehaltung seiner Norm zu einem Funktional auf ganz E fortsetzen, legen wir durch diese Hyperebene in L eine „große" Hyperebene in E, wobei sich der Abstand zum Nullpunkt nicht verkleinert.

Die komplexe Variante des Satzes von HAHN-BANACH (3.2., Satz 4a) liefert ein komplexes Analogon des vorhergehenden Satzes:

Es sei E ein komplexer normierter Raum und f_0 ein beschränktes lineares Funktional auf einem Teilraum $L \subset E$. Dann existiert ein beschränktes lineares Funktional f auf ganz E mit

$$f(x) = f_0(x), \qquad x \in L,$$

$$\|f\|_{\text{auf}\, E} = \|f_0\|_{\text{auf}\, L}.$$

Aus dem Satz von HAHN-BANACH für normierte Räume ergibt sich folgende wichtige Tatsache.

Folgerung. *Es sei E ein normierter Raum. Dann existiert für beliebige Punkte $x_1 \neq x_2$ aus E ein stetiges lineares Funktional f mit $f(x_1) \neq f(x_2)$. Man sagt, daß ein derartiges Funktional die Punkte x_1 und x_2 trennt.*

Diese Behauptung ist gleichbedeutend damit, daß für jedes $x_0 \neq 0$ ein Funktional f existiert, das x_0 und 0 trennt, d. h., für das $f(x_0) \neq 0$ ist. Zur Konstruktion des gesuchten Funktionals betrachten wir zunächst alle Elemente der Form λx_0 und definieren für diese Elemente ein Funktional f_0 durch die Gleichung $f_0(\lambda x_0) = \lambda$. Anschließend setzen wir f_0 (unter Benutzung des Satzes von HAHN-BANACH) auf ganz E fort. Als Ergebnis erhalten wir ein stetiges lineares Funktional f mit $f(x_0) = 1 \neq 0$.

4.1.4. Lineare Funktionale auf abzählbar-normierten Räumen. Es sei E ein abzählbar-normierter Raum mit den Normen $\|\cdot\|_k$ $(k = 1, 2, \ldots)$. Dabei kann man ohne Beschränkung der Allgemeinheit annehmen (vgl. S. 175), daß für jedes $x \in E$

$$\|x\|_1 \leqq \|x\|_2 \leqq \cdots \leqq \|x\|_n \leqq \cdots \tag{9}$$

ist. Es sei f ein stetiges lineares Funktional auf E, dann existiert in E eine Nullumgebung U, auf der f beschränkt ist. Nach Definition der Topologie in einem abzählbar-normierten Raum kann man eine natürliche Zahl k und ein $\varepsilon > 0$ finden, so daß die Kugel $B_{k,\varepsilon} = \{x: \|x\|_k < \varepsilon\}$ ganz in U liegt. Dann ist f auf dieser Kugel beschränkt und damit bezüglich der Norm $\|\cdot\|_k$ beschränkt und stetig. Es existiert also eine Zahl $C > 0$, so daß

$$|f(x)| \leqq C\,\|x\|_k$$

für alle $x \in E$ gilt. Andererseits ist offenbar auch jedes bezüglich einer Norm $\|\cdot\|_n$ beschränkte Funktional stetig auf E. Bezeichnen wir mit E^* die Gesamtheit aller stetigen linearen Funktionale auf E und mit E_n^* die Gesamtheit aller bezüglich der Norm $\|\cdot\|_n$ stetigen linearen Funktionale auf E, so gilt also

$$E^* = \bigcup_{n=1}^{\infty} E_n^*. \tag{10}$$

Dabei folgt aus Bedingung (9), daß

$$E_1^* \subset E_2^* \subset \cdots \subset E_n^* \subset \cdots$$

ist. Wenn f ein stetiges lineares Funktional auf E, d. h. $f \in E^*$ ist, heißt die kleinste Zahl n mit $f \in E_n^*$ die *Ordnung* von f. Nach Gleichung (10) ist also *jedes stetige lineare Funktional auf E von endlicher Ordnung.*

Aufgaben

1. Es ist zu beweisen, daß in einem abzählbar-normierten Raum zu je zwei Elementen $x_1 \neq x_2$ ein stetiges lineares Funktional existiert, das diese Punkte trennt.
2. Dasselbe ist für einen lokalkonvexen Raum zu zeigen.

4.2. Der duale Raum

4.2.1. Definition des dualen Raumes. Für lineare Funktionale kann man eine Addition und eine Skalarmultiplikation erklären. Es seien f_1 und f_2 zwei stetige lineare Funktionale auf einem linearen Raum E. Das Funktional

$$f(x) = f_1(x) + f_2(x), \qquad x \in E,$$

wird dann *Summe* von f_1 und f_2 genannt und mit $f_1 + f_2$ bezeichnet.

Das Funktional

$$f(x) = \alpha f_1(x), \qquad x \in E,$$

heißt *Produkt* von f_1 mit der Zahl α und wird mit αf_1 bezeichnet.

Die Definitionsgleichungen von $f_1 + f_2$ und αf_1 kann man auch wie folgt schreiben:

$$(f_1 + f_2)(x) = f_1(x) + f_2(x), \qquad (\alpha f_1)(x) = \alpha f_1(x).$$

Offenbar sind die Summe $f_1 + f_2$ und das Produkt αf_1 ebenfalls lineare Funktionale. Ist E ein topologischer Raum, so folgt außerdem aus der Stetigkeit der Funktionale f_1 und f_2 auch die Stetigkeit von $f_1 + f_2$ und αf_1.

Man prüft leicht nach, daß die so definierte Addition und Skalarmultiplikation allen Axiomen eines linearen Raumes genügen. Anders ausgedrückt, bildet die Gesamtheit aller stetigen linearen Funktionale auf einem topologischen Raum E einen linearen Raum. Er heißt *dualer* Raum von E und wird mit E^* bezeichnet.

Aufgabe. Die Gesamtheit aller nicht notwendig stetigen linearen Funktionale auf E heißt *algebraisch dualer* Raum und wird mit $E^\#$ bezeichnet. Man gebe einen topologischen linearen Raum E an mit

$$E^* \neq E^\#.$$

Im dualen Raum E^* kann man verschiedene Topologien einführen. Am wichtigsten sind die *starke* und die *schwache* Topologie.

4.2.2. Die starke Topologie im dualen Raum. Wir beginnen mit dem einfachsten Fall, und zwar sei der ursprüngliche Raum E normiert. Für stetige lineare Funktionale auf einem normierten Raum hatten wir den Begriff der Norm eingeführt:

$$\|f\| = \sup_{x \neq 0} \frac{|f(x)|}{\|x\|}.$$

Diese Größe genügt allen Forderungen in der Definition eines normierten Raumes.

Denn es ist

1. $\|f\| > 0$ für jedes nicht identisch verschwindende Funktional f,
2. $\|\alpha f\| = |\alpha| \, \|f\|$,
3. $$\|f_1 + f_2\| = \sup_{x \neq 0} \frac{|f_1(x) + f_2(x)|}{\|x\|}$$
$$\leqq \sup_{x \neq 0} \frac{|f_1(x)|}{\|x\|} + \sup_{x \neq 0} \frac{|f_2(x)|}{\|x\|} = \|f_1\| + \|f_2\|.$$

Im Fall eines normierten Raumes E kann man also den dualen Raum E^* in natürlicher Weise mit der Struktur eines normierten Raumes versehen. Die Topologie in E^*, die der eingeführten Norm entspricht, heißt *starke Topologie* in E^*. Wenn wir hervorheben wollen, daß E^* als normierter Raum betrachtet wird, werden wir $(E^*, \|\cdot\|)$ anstelle von E^* schreiben.

Wir konstatieren folgende wichtige Eigenschaft des dualen Raumes eines normierten Raumes E.

Satz 1. *Der duale Raum* $(E^*, \|\cdot\|)$ *ist vollständig.*

Beweis. Es sei $\{f_n\}$ eine Fundamentalfolge linearer Funktionale. Dann gibt es zu jedem $\varepsilon > 0$ ein N, so daß $\|f_n - f_m\| < \varepsilon$ ist für $n, m \geqq N$. Hieraus erhält man für jedes $x \in E$

$$|f_n(x) - f_m(x)| \leqq \|f_n - f_m\| \cdot \|x\| < \varepsilon \|x\|,$$

d. h. die Zahlenfolge $\{f_n(x)\}$ konvergiert für jedes $x \in E$. Wir setzen

$$f(x) = \lim_{n \to \infty} f_n(x)$$

und werden nachweisen, daß f ein stetiges lineares Funktional darstellt. Die Linearität überprüft man unmittelbar:

$$f(\alpha x + \beta y) = \lim_{n \to \infty} f_n(\alpha x + \beta y)$$
$$= \lim_{n \to \infty} [\alpha f_n(x) + \beta f_n(y)] = \alpha f(x) + \beta f(y).$$

Zum Beweis der Stetigkeit von f gehen wir von der Ungleichung

$$|f_n(x) - f_m(x)| < \varepsilon \|x\|$$

aus. Lassen wir hier $m \to \infty$ streben, so erhalten wir

$$|f(x) - f_n(x)| \leqq \varepsilon \|x\|.$$

Daraus folgt, daß das Funktional $f - f_n$ beschränkt und somit stetig ist. Dann ist aber auch $f = f_n + (f - f_n)$ stetig. Außerdem folgt aus der letzten Ungleichung die Beziehung $\|f - f_n\| \leqq \varepsilon$ für $n \geqq N$, d. h., $\{f_n\}$ konvergiert gegen f. Damit ist der Satz bewiesen.

Wir möchten noch einmal betonen, daß dieser Satz unabhängig davon gültig ist, ob der ursprüngliche Raum vollständig war oder nicht.

Bemerkung. *Der normierte Raum E sei nicht vollständig, mit $\bar{E}$ bezeichnen wir seine Vervollständigung. Dann sind E^* und $(\bar{E})^*$ isomorph.*

Das ergibt sich daraus, daß E eine überall dichte Teilmenge von $\bar{E}$ ist. Daher kann nämlich jedes stetige lineare Funktional f auf E stetig von E auf $\bar{E}$ fortgesetzt werden, und zwar auf genau eine Weise. Wenn wir diese Fortsetzung mit $\bar{f}$ bezeichnen, gilt offenbar $\bar{f} \in (\bar{E})^*$ und $\|\bar{f}\| = \|f\|$. Dabei ist auch jedes Funktional aus $(\bar{E})^*$ die Fortsetzung eines Funktionals aus $(E)^*$ (und zwar die Fortsetzung seiner Einschränkung auf E). Folglich stellt die Abbildung $f \to \bar{f}$ eine isomorphe Abbildung des Raumes E^* auf den Raum $(\bar{E})^*$ dar.

Wir wollen jetzt die *starke Topologie* für den dualen Raum eines beliebigen topologischen linearen Raumes definieren. Im dualen Raum eines normierten Raumes ist die Gesamtheit aller Funktionale mit

$$\|f\| < \varepsilon$$

eine Nullumgebung. Man betrachtet also die Gesamtheit aller Funktionale, für die auf der Einheitskugel in E die Bedingung $|f(x)| < \varepsilon$ gilt, als Nullumgebung im Raum E^*. Nehmen wir dabei alle möglichen ε, so erhalten wir eine Basis von Nullumgebungen. Ist E jetzt ein nicht normierter topologischer linearer Raum, so hat man naturgemäß anstelle der Einheitskugel in E eine beliebige beschränkte Menge A zu nehmen. Die Gesamtheit der linearen Funktionale mit

$$|f(x)| < \varepsilon \quad \text{für alle} \quad x \in A$$

nennt man dann Nullumgebung $U_{\varepsilon,A}$ in E^*. Variieren hier ε und A, so erhalten wir eine Basis von Nullumgebungen in E^*.

Somit wird die *starke Topologie* in E^* durch ein System von Nullumgebungen bestimmt, die von ε und einer beschränkten Menge A abhängen. Daß dieses System von Nullumgebungen den Raum E^* tatsächlich zu einem topologischen linearen Raum macht, wollen wir hier nicht nachprüfen, obwohl das nicht schwer ist (siehe z. B. [9]).

Im Fall eines normierten Raumes E stimmen die mit Hilfe der Norm definierte und die eben beschriebene starke Topologie in E^* offenbar überein.

Wir bemerken noch, daß E^* mit der starken Topologie stets separiert und lokalkonvex ist (unabhängig davon, wie die Topologie in E beschaffen war). Ist nämlich $f_0 \in E^*$ und $f_0 \neq 0$, dann gibt es ein Element $x_0 \in E$ mit $f_0(x_0) \neq 0$. Setzen wir nun $\varepsilon = |f(x_0)|/2$ und $A = \{x_0\}$, so ist $f_0 \notin U_{\varepsilon,A}$, d. h., E^* ist separiert. Daß E^* mit der starken Topologie ein lokalkonvexer Raum ist, ergibt sich einfach daraus, daß die Umgebungen $U_{\varepsilon,A}$ für jedes $\varepsilon > 0$ und jede beschränkte Menge $A \subset E$ in E^* konvex sind. Die starke Topologie in E^* werden wir mit dem Symbol b bezeichnen. Wenn wir besonders hervorheben wollen, daß E^* mit der starken Topologie betrachtet wird, werden wir (E^*, b) anstelle von E^* schreiben.

4.2.3. Beispiele dualer Räume

1. E sei der (reelle oder komplexe) n-dimensionale lineare Raum. Wir wählen in E irgendeine Basis $e_1, \ldots, e_n$. Dann ist jeder Vektor $x \in E$ eindeutig in der Form $x = \sum_{i=1}^{n} x_i e_i$ darstellbar. Wenn nun f ein lineares Funktional auf E ist, gilt offenbar

$$f(x) = \sum_{i=1}^{n} f(e_i)\, x_i. \tag{1}$$

Also wird ein lineares Funktional eindeutig durch seine Werte auf den Basisvektoren $e_1, \ldots, e_n$ bestimmt. Diese Werte können dabei beliebig vorgegeben werden.

Wir definieren nun lineare Funktionale $g_1, \ldots, g_n$, indem wir

$$g_j(e_i) = \begin{cases} 1 & \text{für} \quad i = j, \\ 0 & \text{für} \quad i \neq j \end{cases}$$

setzen. Diese Funktionale sind offensichtlich linear unabhängig, und es gilt $g_j(x) = x_j$. Daher kann man Formel (1) in der Gestalt

$$f(x) = \sum_{i=1}^{n} f(e_i)\, g_i(x)$$

schreiben. Somit bilden die Funktionale $g_1, \ldots, g_n$ eine Basis im Raum E^*, d. h., E^* ist ebenfalls der n-dimensionale lineare Raum. Die Basis $g_1, \ldots, g_n$ in E^* nennt man *dual* zur Basis $e_1, \ldots, e_n$ in E.

Verschiedene Normen in E induzieren verschiedene Normen in E^*. Wir geben einige Beispiele von Paaren einander entsprechender Normen in E und E^* an (und empfehlen dem Leser, die zugehörigen Beweise sorgfältig durchzuführen):

a) $\|x\| = \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2}, \qquad \|f\| = \left(\sum_{i=1}^{n} |f_i|^2\right)^{1/2};$

b) $\|x\| = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}, \qquad \|f\| = \left(\sum_{i=1}^{n} |f_i|^q\right)^{1/q},$

$\frac{1}{p} + \frac{1}{q} = 1, \qquad 1 < p < \infty;$

c) $\|x\| = \sup_{1 \leq i \leq n} |x_i|, \qquad \|f\| = \sum_{i=1}^{n} |f_i|;$

d) $\|x\| = \sum_{i=1}^{n} |x_i|, \qquad \|f\| = \sup_{1 \leq i \leq n} |f_i|.$

In diesen Formeln sind $x_1, \ldots, x_n$ die Koordinaten eines Vektors $x \in E$ bezüglich der Basis $e_1, \ldots, e_n$ und $f_1, \ldots, f_n$ die Koordinaten eines Funktionals $f \in E^*$ bezüglich der dualen Basis $g_1, \ldots, g_n$.

Aufgabe. Man beweise, daß alle aufgezählten Normen im n-dimensionalen Raum ein und dieselbe Topologie definieren.

2. Wir betrachten den Raum c_0 aller Nullfolgen $x = (x_1, x_2, \ldots, x_n, \ldots)$ mit der Norm $\|x\| = \sup\limits_n |x_n|$. Dann werden wir zeigen, daß *der dazu duale Raum* $(c_0^*, \|\cdot\|)$ *dem Raum* l_1 *aller absolut-summierbaren Folgen* $f = (f_1, f_2, \ldots, f_n, \ldots)$ *mit der Norm* $\|f\| = \sum\limits_{n=1}^{\infty} |f_n|$ *isometrisch-isomorph ist.*

Eine beliebige Folge $f \in l_1$ definiert im Raum c_0 ein beschränktes lineares Funktional $\tilde{f}$ durch die Formel

$$\tilde{f}(x) = \sum_{n=1}^{\infty} f_n x_n. \tag{2}$$

Wegen $|\tilde{f}(x)| \leqq \|x\| \sum\limits_{n=1}^{\infty} |f_n|$ gilt dabei $\|\tilde{f}\| \leqq \sum\limits_{n=1}^{\infty} |f_n| \leqq \|f\|$. Um auch die umgekehrte Abschätzung zu erhalten, betrachten wir in c_0 die Vektoren

$$e_1 = (1, 0, 0, \ldots, 0, 0, \ldots),$$
$$e_2 = (0, 1, 0, \ldots, 0, 0, \ldots),$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$e_n = (0, 0, 0, \ldots, 1, 0, \ldots),$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

und setzen damit $x^{(N)} = \sum\limits_{n=1}^{N} \frac{f_n}{|f_n|} e_n$ (im Fall $f_n = 0$ setzen wir $\frac{f_n}{|f_n|} = 0$). Dann ist $x^{(N)} \in c_0$, $\|x^{(N)}\| \leqq 1$ und

$$\tilde{f}(x^{(N)}) = \sum_{n=1}^{N} \frac{f_n}{|f_n|} \tilde{f}(e_n) = \sum_{n=1}^{N} |f_n|,$$

woraus sich $\lim\limits_{N\to\infty} \tilde{f}(x^{(N)}) = \sum\limits_{n=1}^{\infty} |f_n| = \|f\|$ ergibt. Also ist $\|\tilde{f}\| \geqq \sum\limits_{n=1}^{\infty} |f_n|$, und zusammen mit der oben gezeigten entgegengesetzten Ungleichung folgt

$$\|\tilde{f}\| = \sum_{n=1}^{\infty} |f_n| = \|f\|.$$

Auf diese Weise haben wir eine isometrische lineare Abbildung $f \to \tilde{f}$ des Raumes l_1 in den Raum c_0^* erhalten. Es muß noch gezeigt werden, daß das Bild von l_1 bei dieser Abbildung mit ganz c_0^* übereinstimmt, d. h., daß jedes Funktional $\tilde{f} \in c_0^*$ in der Form (2) mit $f = \{f_n\} \in l_1$ darstellbar ist. Dazu schreiben wir jedes $x = \{x_n\} \in c_0$ in der Form $x = \sum\limits_{n=1}^{\infty} x_n e_n$; die rechts stehende Reihe konvergiert

dabei wegen $\left\| x - \sum_{n=1}^{N} x_n e_n \right\| = \sup_{n>N} |x_n| \to 0$ für $N \to \infty$ gegen das Element x. Da das Funktional $\tilde{f} \in c_0^*$ stetig ist, gilt $\tilde{f}(x) = \sum_{n=1}^{\infty} x_n \tilde{f}(e_n)$. Es muß also nur noch die Bedingung $\sum_{n=1}^{\infty} |\tilde{f}(e_n)| < \infty$ nachgewiesen werden. Wenn wir $x^{(N)} = \sum_{n=1}^{N} \frac{\tilde{f}(e_n)}{|\tilde{f}(e_n)|} e_n$ setzen, ergibt sich wegen $x^{(N)} \in c_0$ und $\|x^{(N)}\| \leqq 1$

$$\sum_{n=1}^{N} |\tilde{f}(e_n)| = \sum_{n=1}^{N} \frac{\tilde{f}(e_n)}{|\tilde{f}(e_n)|} \tilde{f}(e_n) = \tilde{f}(x^{(N)}) \leqq \|\tilde{f}\|.$$

Daraus folgt $\sum_{n=1}^{\infty} |\tilde{f}(e_n)| < \infty$, denn N war beliebig.

3. Es ist nicht schwer zu zeigen, daß der zu l_1 duale Raum l_1^* dem Raum m der beschränkten Folgen $x = \{x_n\}$ mit der Norm $\|x\| = \sup_n |x_n|$ isometrisch-isomorph ist.

4. Es sei $p > 1$ und l_p der Raum aller Folgen $x = \{x_n\}$ mit

$$\|x\| = \left(\sum_{n=1}^{\infty} |x_n|^p \right)^{1/p} < \infty.$$

Man kann beweisen, daß der dazu duale Raum l_p^* dem Raum l_q, $\frac{1}{p} + \frac{1}{q} = 1$, isometrisch-isomorph ist. Jedes lineare Funktional auf l_p ist dabei von der Form

$$f(x) = \sum_{n=1}^{\infty} f_n x_n, \qquad x = \{x_n\} \in l_p, \qquad f = \{f_n\} \in l_q.$$

Der Beweis beruht auf der Anwendung der Hölderschen Ungleichung.

5. Wir wollen jetzt die Struktur des dualen Raumes eines Hilbertraumes klären.

Satz 2. *H sei ein reeller Hilbertraum. Jedes stetige lineare Funktional f auf H läßt sich mit einem eindeutig bestimmten $x_0 \in H$ in der Form*

$$f(x) = (x, x_0), \qquad x \in H, \tag{3}$$

darstellen, wobei $\|f\| = \|x_0\|$ ist. Umgekehrt definiert Formel (3) *für jedes $x_0 \in H$ ein stetiges lineares Funktional f mit $\|f\| = \|x_0\|$. Somit definiert Gleichung* (3) *einen isometrischen Isomorphismus $f \to x_0$ zwischen den Räumen H^* und H.*

Beweis. Offenbar definiert Formel (3) für jedes $x_0 \in H$ ein lineares Funktional auf H. Wegen $|f(x)| = |(x, x_0)| \leqq \|x\| \|x_0\|$ ist dieses Funktional stetig, und wegen $f(x_0) = \|x_0\|^2$ gilt $\|f\| = \|x_0\|$.

Wir wollen zeigen, daß auch jedes stetige lineare Funktional f auf H in der Form (3) darstellbar ist. Wenn $f = 0$ ist, können wir $x_0 = 0$ setzen. Es sei also $f \neq 0$ und $H_0 = \{x : f(x) = 0\}$ der Kern des Funktionals f. Wegen der Stetigkeit von f ist H_0

ein *abgeschlossener* linearer Teilraum von H. In 3.1.6. wurde gezeigt, daß die Codimension des Kerns eines beliebigen linearen Funktionals gleich 1 ist. Deshalb folgt nach 3.1.6., Folgerung 3, daß das orthogonale Komplement $H_0^{\perp}$ des Teilraumes H_0 eindimensional ist. Also existiert ein (von Null verschiedener) zu H_0 orthogonaler Vektor y_0, so daß jeder Vektor $x \in H$ eindeutig in der Form $x = y + \lambda y_0$ mit $y \in H_0$ darstellbar ist. Offenbar kann man dabei annehmen, daß $\|y_0\| = 1$ ist. Setzen wir nun $x_0 = f(y_0)\, y_0$, so gilt für jedes $x \in H$

$$x = y + \lambda y_0, \qquad y \in H_0,$$

$$f(x) = \lambda f(y_0),$$

$$(x, x_0) = \lambda(y_0, x_0) = \lambda f(y_0)\,(y_0, y_0) = \lambda f(y_0).$$

Damit ist $f(x) = (x, x_0)$ für alle $x \in H$. Dabei ist x_0 eindeutig bestimmt. Ist nämlich $f(x) = (x, x_0')$, $x \in H$, so folgt $(x, x_0 - x_0') = 0$ für alle $x \in H$. Setzt man hier $x = x_0 - x_0'$, so ergibt sich $x_0 = x_0'$. Damit ist der Satz bewiesen.

Bemerkungen

1. E sei ein nicht vollständiger unitärer Raum und der Hilbertraum H seine Vervollständigung. Da die Räume E^* und H^* isometrisch-isomorph sind (vgl. die Bemerkung auf S. 185) und H^* isometrisch-isomorph zu H ist, gilt die folgende Behauptung: *Der duale Raum E^* eines nicht vollständigen unitären Raumes E ist der Vervollständigung H von E isometrisch-isomorph.*

2. Satz 2 gilt auch für komplexe Hilberträume (der Beweis ist vollkommen derselbe, wenn man $x_0 = f(y_0)\, y_0$ durch $x_0 = \overline{f(y_0)}\, y_0$ ersetzt). Es besteht nur ein Unterschied zwischen dem komplexen und dem reellen Fall. Und zwar ist die Abbildung, die dem Element $x_0 \in H$ das Funktional $f(x) = (x, x_0)$ zuordnet, jetzt ein antilinearer isometrischer Isomorphismus, d. h., dem Element λx_0 entspricht das Element $\bar{\lambda} f$.

6. In den Beispielen 1 bis 5 wurden normierte Räume betrachtet. Jetzt wollen wir einen abzählbar-normierten Raum untersuchen. Φ sei der Raum aller Folgen $x = \{x_n\}$ von reellen Zahlen, für die

$$\|x\|_k = \left(\sum_{n=1}^{\infty} n^k x_n^2\right)^{1/2} < \infty, \qquad k = 1, 2, \ldots,$$

ist. Diese Normen ergeben sich aus den Skalarprodukten

$$(x, y)_k = \sum_{n=1}^{\infty} n^k x_n y_n, \qquad k = 1, 2, \ldots$$

Der Raum Φ ist mit dem Skalarprodukt $(\cdot, \cdot)_k$ ein unitärer Raum, Φ_k sei dessen Vervollständigung. Es ist leicht zu sehen, daß man Φ_k und den Hilbertraum aller Folgen $x = \{x_n\}$ mit $\|x\|_k < \infty$ identifizieren kann. Nach Satz 2 sind Φ_k und sein dualer Raum Φ_k^* isometrisch-isomorph. Bei diesem Isomorphismus wird jedem stetigen

linearen Funktional $f \in \Phi_k^*$ eine Folge $\tilde{f} = \{f_n\}$ mit

$$\|f\| = \left(\sum_{n=1}^{\infty} n^k |f_n|^2\right)^{1/2} < \infty,$$

$$f(x) = (x, \tilde{f})_k = \sum_{n=1}^{\infty} n^k x_n f_n, \qquad x = \{x_n\} \in \Phi_k$$

zugeordnet, und umgekehrt bestimmt jede solche Folge ein Element aus Φ_k^*. Definieren wir hier das Funktional $f \in \Phi_k^*$ nicht mit Hilfe der Folge $\{f_n\}$, sondern durch die Folge $\{g_n\}$, $g_n = n^k f_n$, so ist

$$f(x) = \sum_{n=1}^{\infty} x_n g_n \quad \text{und} \quad \|f\| = \left(\sum_{n=1}^{\infty} n^{-k} g_n^2\right)^{1/2}.$$

Daher kann man Φ_k^* und den Hilbertraum aller Folgen $\{g_n\}$ mit

$$\sum_{n=1}^{\infty} n^{-k} g_n^2 < \infty, \tag{4}$$

in dem das Skalarprodukt durch die Formel

$$(g^{(1)}, g^{(2)}) = \sum_{n=1}^{\infty} n^{-k} g_n^{(1)} g_n^{(2)}$$

erklärt ist, identifizieren. Wegen $\Phi^* = \bigcup_{k=1}^{\infty} \Phi_k^*$ ist Φ^* der Raum aller Folgen $\{g_n\}$ für die ein k existiert, so daß die Bedingung (4) erfüllt ist. Der Wert eines solchen Funktionals $\{g_n\}$ wird für jedes Element $x = \{x_n\} \in \Phi$ durch den Ausdruck $\sum_{n=1}^{\infty} x_n g_n$ gegeben.

Ist also der Raum Φ Durchschnitt einer fallenden Kette von Hilberträumen

$$\Phi_1 \supset \Phi_2 \supset \cdots \supset \Phi_k \supset \cdots, \qquad \Phi = \bigcap_{k=1}^{\infty} \Phi_k,$$

dann ist Φ^* die Vereinigung der wachsenden Kette der Hilberträume

$$\Phi_1^* \subset \Phi_2^* \subset \cdots \subset \Phi_k^* \subset \cdots, \qquad \Phi^* = \bigcup_{k=1}^{\infty} \Phi_k^*.$$

Hier ist es zweckmäßig, die Bezeichnung $\Phi_k^* = \Phi_{-k}$ einzuführen und den Raum l_2 mit Φ_0 zu bezeichnen. Dann erhalten wir eine in beiden Richtungen unendliche Kette von Hilberträumen

$$\cdots \subset \Phi_k \subset \cdots \subset \Phi_1 \subset \Phi_0 \subset \Phi_{-1} \subset \cdots \subset \Phi_{-k} \subset \cdots,$$

in der $\Phi_k^* = \Phi_{-k}$ für jedes $k = 0, \pm 1, \pm 2, \ldots$ ist.

4.2.4. Der biduale Raum. Die stetigen linearen Funktionale auf einem topologischen linearen Raum E bilden selbst wieder einen topologischen linearen Raum, den zu E dualen Raum (E^*, b). Somit kann man auch vom Raum E^{**} aller stetigen linearen Funktionale auf E^* sprechen, der *bidualer Raum* von E genannt wird.

Jedes Element x_0 aus E definiert ein stetiges lineares Funktional auf E^*, und zwar durch die Gleichung

$$\psi_{x_0}(f) = f(x_0), \tag{5}$$

wobei f ganz E^* durchläuft. Durch Gleichung (5) wird jedem $f \in E^*$ eine Zahl $\psi_{x_0}(f)$ zugeordnet, d. h. ein Funktional ψ_{x_0} auf E^* erklärt. Wegen

$$\psi_{x_0}(\alpha f_1 + \beta f_2) = \alpha f_1(x_0) + \beta f_2(x_0) = \alpha \psi_{x_0}(f_1) + \beta \psi_{x_0}(f_2)$$

ist dieses Funktional linear.

Jedes solche Funktional ist auch stetig auf E^*. Zum Beweis sei $\varepsilon > 0$ beliebig. Dann betrachten wir in E^* die Nullumgebung $U_{\varepsilon,A}$, wobei A eine beschränkte Menge ist, die den Punkt x_0 enthält. Nach Definition von $U_{\varepsilon,A}$ gilt

$$|\psi_{x_0}(f)| = |f(x_0)| < \varepsilon \quad \text{für} \quad f \in U_{\varepsilon,A}.$$

Das bedeutet aber, daß das Funktional ψ_{x_0} im Nullpunkt und somit im ganzen Raum E^* stetig ist.

Also haben wir eine Abbildung des Raumes E auf eine gewisse Teilmenge des Raumes E^{**} erhalten. Diese offenbar lineare Abbildung von E in E^{**} wird *kanonische Abbildung des Raumes E in seinen bidualen Raum* genannt und mit π bezeichnet. Wenn es hinreichend viele lineare Funktionale auf E gibt (wenn E zum Beispiel normiert oder wenigstens lokalkonvex und separiert ist), ist diese Abbildung umkehrbar eindeutig. Denn dann existiert zu je zwei verschiedenen Punkten $x', x'' \in E$ ein Funktional $f \in E^*$ mit $f(x') \neq f(x'')$, so daß $\psi_{x'}$ und $\psi_{x''}$ verschiedene Funktionale auf E^* sind. Gilt darüber hinaus $\pi(E) = E^{**}$, so heißt der (separierte lokalkonvexe) Raum E *halbreflexiv*. Im Raum E^{**} (dem zu (E^*, b) dualen Raum) kann man die starke Topologie einführen, die wir mit b^* bezeichnen wollen. Ist der Raum E halbreflexiv und die Abbildung $\pi: E \to E^{**}$ stetig, dann heißt E *reflexiv*. Man kann zeigen, daß die Abbildung π^{-1} immer stetig ist. *Daher stellt die kanonische Abbildung $\pi: E \to E^{**}$ im Fall eines reflexiven Raumes E einen Isomorphismus zwischen den topologischen linearen Räumen E und $E^{**} = (E^{**}, b^*)$ dar.*

Da jetzt jedes Element aus E auch als Element des Raumes E^{**} angesehen werden kann, verwendet man für die Werte des Funktionals $f \in E^*$ anstelle der Schreibweise $f(x)$ günstiger die symmetrische Bezeichnung

$$f(x) = (f, x). \tag{6}$$

Bei festem $f \in E^*$ können wir (f, x) dann als Funktional auf E und bei festem x als Funktional auf E^* (wobei x als Element aus E^{**} interpretiert wird) auffassen.

Ist E normiert (und sind somit auch die Räume E^*, E^{**} normiert), *so ist die kanonische Abbildung von E in E^{**} ein isometrischer Isomorphismus.*

Zum Beweis betrachten wir ein beliebiges Element x aus E mit der Norm $\|x\|$. Wenn wir die Norm seines Bildes in E^{**} mit $\|x\|_2$ bezeichnen, müssen wir die Gleichheit $\|x\| = \|x\|_2$ zeigen. Ist f ein beliebiges von Null verschiedenes Element aus E^*, so gilt

$$|(f, x)| \leqq \|f\| \cdot \|x\|, \quad \text{d. h.} \quad \|x\| \geqq \frac{|(f, x)|}{\|f\|}.$$

Da die linke Seite der letzten Ungleichung nicht von f abhängt, ergibt sich

$$\|x\| \geqq \sup \frac{|(f, x)|}{\|f\|} = \|x\|_2.$$

Andererseits findet sich nach dem Satz von HAHN-BANACH zu jedem $x_0 \in E$ ein von Null verschiedenes lineares Funktional f_0 mit

$$|(f_0, x_0)| = \|f_0\| \cdot \|x_0\|. \tag{7}$$

(Zur Konstruktion eines solchen Funktionals kann man etwa $f_0(x) = \alpha \neq 0$ für alle Elemente der Form $x = \alpha x_0$ setzen und dieses Funktional unter Beibehaltung der Norm auf ganz E fortsetzen.) Aus (7) folgt dann

$$\|x\|_2 = \sup_{f \in E^*} \frac{|(f, x)|}{\|f\|} \geqq \|x\|,$$

d. h. $\|x\| = \|x\|_2$, was zu beweisen war. Somit ist ein normierter Raum E isometrisch zur linearen Menge $\pi(E)$ in E^{**} (die im allgemeinen nicht vollständig ist). Wenn man E mit $\pi(E)$ identifiziert, kann man auch sagen, daß $E \subset E^{**}$ ist.

Da die kanonische Abbildung $\pi: E \to E^{**}$ für normierte Räume eine Isometrie ist, *stimmen die Begriffe der Halbreflexivität und Reflexivität für normierte Räume überein.*

Weil der duale Raum eines normierten Raumes vollständig ist, *muß jeder reflexive normierte Raum E vollständig sein.*

Die endlichdimensionalen euklidischen Räume und die Hilberträume stellen einfachste Beispiele reflexiver Räume dar (für sie gilt sogar $E = E^*$).

Der Raum c_0 aller Nullfolgen ist ein Beispiel eines vollständigen nicht reflexiven Raumes. Wie wir oben gezeigt haben (4.2., Beispiel 2), ist der Raum l_1 aller absolut konvergenten Reihen dual zu c_0, während der Raum m aller beschränkten Folgen dual zu l_1 ist.

Der Raum $C[a, b]$ der stetigen Funktionen auf dem Intervall $[a, b]$ ist ebenfalls nicht reflexiv. Wir werden jedoch diese Behauptung hier nicht beweisen.[1])

[1]) Man kann sogar die folgende schärfere Aussage beweisen: Es gibt keinen normierten Raum, zu dem $C[a, b]$ der duale Raum ist.

Als Beispiel eines reflexiven Raumes, der nicht mit seinem dualen Raum übereinstimmt, kann der Raum l_p für $1 < p$, $p \neq 2$, dienen (hier gilt $l_p^* = l_q$, wobei $1/p + 1/q = 1$ ist, und $l_p^{**} = l_q^* = l_p$).

Aufgabe. Man beweise, daß ein abgeschlossener Teilraum eines reflexiven Raumes reflexiv ist.

4.3. Die schwache Topologie und die schwache Konvergenz

4.3.1. Die schwache Topologie und die schwache Konvergenz in einem topologischen linearen Raum. Wir betrachten einen topologischen linearen Raum E und die Gesamtheit aller stetigen Funktionale auf E. Sind $f_1, \ldots, f_n$ endlich viele stetige Funktionale und ist ε eine positive Zahl, so ist die Menge

$$\{x: |f_i(x)| < \varepsilon;\ i = 1, 2, \ldots, n\} \tag{1}$$

offen in E und enthält den Nullpunkt, d. h., sie ist eine Nullumgebung. Der Durchschnitt zweier solcher Umgebungen enthält stets eine Menge der Form (1), und daher bilden alle Mengen der Form (1) eine Basis von Nullumgebungen für eine gewisse Topologie in E. Diese Topologie heißt *schwache Topologie* im Raum E. Die schwache Topologie in E ist die schwächste Topologie, in der noch alle linearen Funktionale stetig sind, die in der ursprünglichen Topologie dieses Raumes stetig waren.

Offenbar ist jede im Sinne der schwachen Topologie offene Menge auch in der ursprüngliche Topologie des Raumes E offen. Jedoch ist die Umgekehrung allgemein nicht richtig (die Mengen der Form (1) brauchen keine Basis von Nullumgebungen für die ursprüngliche Topologie zu bilden). In der in 2.5. verwendeten Terminologie bedeutet das, daß die schwache Topologie des Raumes E schwächer als die ursprüngliche Topologie ist. Dadurch wird die für sie verwendete Bezeichnung gerechtfertigt.

Wenn in E hinreichend viele stetige lineare Funktionale existieren (z. B. wenn E normiert ist), dann erfüllt die schwache Topologie in E das Hausdorffsche Trennungsaxiom. Man kann auch leicht nachweisen, daß die Addition und die Skalarmultiplikation in E bezüglich der schwachen Topologie dieses Raumes stetig sind.

Sogar im Fall eines normierten Raumes braucht die schwache Topologie in E nicht dem ersten Abzählbarkeitsaxiom zu genügen. Folglich läßt sich diese Topologie im allgemeinen nicht mit Hilfe von konvergenten Folgen beschreiben. Trotzdem stellt die Konvergenz, die durch diese Topologie definiert wird, einen wichtigen Begriff dar. Man nennt sie *schwache Konvergenz*. Im Unterschied dazu heißt die Konvergenz, die durch die ursprüngliche Topologie des Raumes E definiert wird (die Normkonvergenz, falls E normiert ist), *starke Konvergenz*.

Der Begriff der schwachen Konvergenz läßt sich auch folgendermaßen formulieren: Eine Folge $\{x_n\}$ von Elementen aus E heißt *schwach konvergent* gegen $x_0 \in E$, wenn für jedes stetige lineare Funktional $\varphi(x)$ auf E die Zahlenfolge $\{\varphi(x_n)\}$ gegen $\varphi(x_0)$ konvergiert.

Zum Beweis nehmen wir der Einfachheit halber $x_0 = 0$ an, und wir setzen $\varphi(x_n) \to 0$ für jedes $\varphi \in E^*$ voraus. Dann gibt es zu jeder schwachen Nullumgebung

$$U = \{x: |\varphi_i(x)| < \varepsilon, \ i = 1, 2, \ldots, k\}$$

ein N mit $x_n \in U$ für alle $n \geqq N$ (dazu braucht man nur N_i so zu wählen, daß $|\varphi_i(x_n)| < \varepsilon$ für $n \geqq N_i$ ist, und anschließend $N = \max N_i$ zu setzen). Wenn umgekehrt zu jeder schwachen Nullumgebung U ein N mit $x_n \in U$ für alle $n \geqq N$ existiert, ist die Beziehung $\lim\limits_{n\to\infty} \varphi(x_n) = 0$ offenbar für jedes feste $\varphi \in E^*$ erfüllt.

Da jede schwache Topologie eines Raumes E schwächer als dessen starke Topologie ist, *konvergiert jede stark konvergente Folge auch schwach.* Das Umgekehrte ist allgemein nicht richtig (siehe Beispiele unten).

4.3.2. Die schwache Konvergenz in normierten Räumen. Wir wollen den Begriff der schwachen Konvergenz für normierte Räume ausführlicher untersuchen.

Satz 1. *Zu jeder schwach konvergenten Folge $\{x_n\}$ in einem normierten Raum existiert eine Konstante C mit*

$$\|x_n\| \leqq C.$$

Anders ausgedrückt: *Jede schwach konvergente Folge in einem normierten Raum ist beschränkt.*

Beweis. Ist die Folge $\{x_n\}$ auf einer Kugel $S[f_0, \varepsilon]$ beschränkt, so ist sie auch auf der Kugel $S[0, \varepsilon] = \{g: \|g\| \leqq \varepsilon\}$ beschränkt. Denn für jedes $g \in S[0, \varepsilon]$ gilt $f_0 + g \in S[f_0, \varepsilon]$ und $(g, x_n) = (f_0 + g, x_n) - (f_0, x_n)$. Dabei ist die Zahlenfolge (f_0, x_n) wegen der schwachen Konvergenz der Folge $\{x_n\}$ beschränkt. Wenn aber $|(g, x_n)| \leqq C$ für alle $g \in S[0, \varepsilon]$ ist, ergibt sich aus der isometrischen Einbettung von E in E^{**}

$$\|x_n\| \leqq \frac{C}{\varepsilon},$$

d. h., die Normen $\|x_n\|$ besitzen eine gemeinsame obere Schranke. Ist nun die Folge $\{x_n\}$ nicht beschränkt, so ist sie auch auf jeder Kugel in E^* nicht beschränkt. Wenn wir also eine Kugel $B_0 \subset E^*$ hernehmen, gibt es einen Index n_1 und ein Element $f \in B_0$ mit $|(f, x_{n_1})| > 1$. Auf Grund der Stetigkeit von (f, x) in Abhängigkeit von f ist die Ungleichung $|(f, x_{n_1})| > 1$ sogar für alle f aus einer abgeschlossenen Kugel $B_1 \subset B_0$ erfüllt. Analog können wir einen Index n_2 und in B_1 eine abgeschlossene Kugel B_2 finden, so daß für alle $f \in B_2$ die Ungleichung $|(f, x_{n_2})| > 2$ erfüllt ist, usw. Allgemein können wir zu jedem k einen Index n_k und eine Kugel $B_k \subset B_{k-1}$ finden, so daß

$$|(f, x_{n_k})| > k \quad \text{für} \quad f \in B_k$$

ist. Dabei kann man annehmen, daß die Radien der Kugeln B_k für $k \to \infty$ gegen Null streben. Da der Raum E^* vollständig ist, existiert ein Element $\tilde{f}$, das zu allen B_k gehört (Cantorscher Durchschnittssatz!). Dann gilt aber

$$|(\tilde{f}, x_{n_k})| > k$$

für alle k, was der schwachen Konvergenz der Folge $\{x_n\}$ widerspricht.

Bemerkung. Beim Beweis der Normbeschränktheit der Folge $\{x_n\}$ haben wir nur ausgenutzt, daß die Zahlenfolge (f, x_n) für jedes $f \in E^*$ beschränkt ist. Somit existiert zu einer Folge $\{x_n\}$, für die alle Zahlenfolgen (f, x_n), $f \in E^*$ beschränkt sind, eine Konstante C mit $\|x_n\| \leqq C$. Diese Aussage läßt sich verallgemeinern: *Jede schwach beschränkte* (d. h. in der schwachen Topologie beschränkte) *Teilmenge Q eines normierten Raumes E ist stark beschränkt* (d. h. in einer hinreichend großen Kugel enthalten). Um das zu zeigen, nehmen wir an, daß es eine Folge $\{x_n\} \subset Q$ mit $\|x_n\| \to \infty$ für $n \to \infty$ gibt. Da Q schwach beschränkt ist, muß auch die Menge $\{x_n\}$ schwach beschränkt sein. Das bedeutet aber, daß diese Menge von jeder schwachen Nullumgebung absorbiert wird. Insbesondere gibt es also zu jedem $f \in E^*$ ein N mit $\{x_n\} \subset N\{x: |(f, x)| < 1\}$, d. h. mit $|(f, x_n)| < N$ für alle n. Nach der obigen Bemerkung widerspricht das jedoch der Voraussetzung $\|x_n\| \to \infty$.

Die schwache Beschränktheit einer Menge Q bedeutet, daß jedes stetige lineare Funktional auf Q beschränkt ist. Damit erhalten wir folgendes wichtige Resultat: *Eine Teilmenge Q eines normierten Raumes ist genau dann beschränkt, wenn jedes Funktional $f \in E^*$ auf Q beschränkt ist.*

In konkreten Fällen ist der folgende Satz zum Nachweis der schwachen Konvergenz einer Folge recht nützlich.

Satz 2. *Eine Folge $\{x_n\}$ von Elementen eines normierten Raumes E ist schwach konvergent gegen $x \in E$, wenn*

1. *die Zahlen $\|x_n\|$ durch eine gemeinsame Konstante M beschränkt sind,*

2. *$f(x_n) \to f(x)$ gilt für jedes f aus einer Menge Δ, deren lineare Hülle überall dicht in E^* ist.*

Beweis. Nach Bedingung 2 und nach Definition der Operationen zwischen linearen Funktionalen gilt

$$\varphi(x_n) \to \varphi(x)$$

für jede Linearkombination φ von Elementen aus Δ. Nun sei φ ein beliebiges Element aus E^* und $\{\varphi_k\}$ eine gegen φ konvergente Folge von Linearkombinationen von Elementen aus Δ. Dann wollen wir zeigen, daß $\varphi(x_n) \to \varphi(x)$ strebt.

Dazu wählen wir ein M mit

$$\|x_n\| \leqq M \quad (n = 1, 2, \ldots) \qquad \text{und} \qquad \|x\| \leqq M$$

und schätzen die Differenz $|\varphi(x_n) - \varphi(x)|$ ab. Wegen $\varphi_k \to \varphi$ existiert zu jedem $\varepsilon > 0$ ein K, so daß $\|\varphi - \varphi_k\| < \varepsilon$ ist für $k \geqq K$. Daher gilt

$$\begin{aligned} |\varphi(x_n) - \varphi(x)| &\leqq |\varphi(x_n) - \varphi_k(x_n)| + |\varphi_k(x_n) - \varphi_k(x)| + |\varphi_k(x) - \varphi(x)| \\ &\leqq \varepsilon M + \varepsilon M + |\varphi_k(x_n) - \varphi_k(x)|. \end{aligned}$$

Nach Bedingung 2 strebt hier $\varphi_k(x_n) \to \varphi_k(x)$ für $n \to \infty$. Folglich strebt $\varphi(x_n) \to \varphi(x)$ für $n \to \infty$, und zwar für jedes $\varphi \in E^*$.

Beispiele. Wir wollen sehen, welche Bedeutung der Begriff der schwachen Konvergenz in konkreten Räumen besitzt.

1. *Im endlichdimensionalen euklidischen Raum* $\boldsymbol{R}^n$ stimmt die schwache Konvergenz mit der starken überein. Um das zu zeigen, betrachten wir eine orthonormierte Basis $e_1, \ldots, e_n$ im $\boldsymbol{R}^n$ und eine Folge $\{x_k\}$, die schwach gegen ein Element x konvergiert. Es sei

$$x_k = x_k^{(1)} e_1 + \cdots + x_k^{(n)} e_n$$

und

$$x = x^{(1)} e_1 + \cdots + x^{(n)} e_n .$$

Dann folgt

$$x_k^{(1)} = (x_k, e_1) \to (x, e_1) = x^{(1)},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$x_k^{(n)} = (x_k, e_n) \to (x, e_n) = x^{(n)},$$

d. h., die Folge $\{x_n\}$ konvergiert koordinatenweise gegen x. Damit gilt aber auch

$$\varrho(x_k, x) = \left(\sum_{i=1}^{n} (x_k^i - x^i)^2 \right)^{1/2} \to 0,$$

d. h., die Folge $\{x_n\}$ ist stark konvergent gegen x. Da aus der starken Konvergenz stets die schwache Konvergenz folgt, ist die Gleichwertigkeit der beiden Konvergenzbegriffe im $\boldsymbol{R}^n$ bewiesen.

2. *Die schwache Konvergenz in* l_2. Hinreichend für die schwache Konvergenz einer beschränkten Folge $\{x_k\}$ gegen ein Element x sind hier die Bedingungen

$$(x_k, e_i) = x_k^{(i)} \to x^{(i)} = (x, e_i), \qquad i = 1, 2, \ldots;$$

dabei ist

$$e_1 = (1, 0, 0, \ldots), \; e_2 = (0, 1, 0, \ldots), \ldots$$

Denn die Linearkombinationen der Elemente e_i sind überall dicht im Raum l_2 (der, wie wir sahen, mit seinem dualen Raum übereinstimmt), und daher ergibt sich die Behauptung aus Satz 2.

Somit bedeutet die schwache Konvergenz einer beschränkten Folge $\{x_k\}$ in l_2, daß die Zahlenfolge $x_k^{(i)}$ der Koordinaten dieser Vektoren für jedes $i = 1, 2, \ldots$ konvergiert. Anders ausgedrückt: Die schwache Konvergenz stimmt mit der koordinatenweisen Konvergenz überein (unter der Bedingung der Beschränktheit).

Unschwer ist zu sehen, daß in l_2 die schwache Konvergenz nicht mit der starken Konvergenz übereinstimmt. Dazu zeigen wir, daß die Folge $e_1, e_2, \ldots, e_n, \ldots$ in l_2 schwach gegen Null konvergiert. Jedes stetige lineare Funktional f auf l_2 läßt sich

als Skalarprodukt $f(x) = (x, a)$ der Vektoren $x \in l_2$ mit einem festen Vektor $a = (a_1, a_2, \ldots)$ aus l_2 schreiben, daher ist

$$f(e_n) = a_n .$$

Für jedes $a \in l_2$ gilt nun $a_n \to 0$ für $n \to \infty$. Daraus folgt

$$\lim_{n \to \infty} f(e_n) = 0$$

für jedes stetige lineare Funktional auf l_2. Im Sinne der starken Topologie besitzt die Folge $\{e_n\}$ jedoch keinen Grenzwert.

Aufgaben

1. In einem Hilbertraum H konvergiere die Folge $\{x_n\}$ schwach gegen x; dabei gelte $\|x_n\| \to \|x\|$ für $n \to \infty$. Es ist zu beweisen, daß die Folge $\{x_n\}$ dann auch stark gegen x konvergiert, d. h., daß $\|x_0 - x\| \to 0$ strebt.

2. Man beweise, daß die Behauptung von Aufgabe 1 auch dann noch gilt, wenn die Bedingung $\|x_n\| \to \|x\|$ durch die Bedingung $\|x_n\| \leqq \|x\|$ für alle n oder die Bedingung $\lim\limits_{n\to\infty} \|x_n\| \leqq \|x\|$ ersetzt wird.

3. H sei ein (separabler) Hilbertraum und Q eine beschränkte Teilmenge von H. Dann kann die Topologie in Q, die durch die schwache Topologie des Raumes H induziert wird, mit Hilfe einer Metrik definiert werden.

4. Man zeige, daß jede abgeschlossene konvexe Teilmenge eines Hilbertraumes in der schwachen Topologie abgeschlossen ist. (Insbesondere ist also jeder abgeschlossene lineare Teilraum eines Hilbertraumes schwach abgeschlossen.) Man gebe ein Beispiel einer abgeschlossenen Menge in einem Hilbertraum an, die nicht schwach abgeschlossen ist.

3. *Die schwache Konvergenz im Raum $C[a, b]$ der stetigen Funktionen.* Die Folge $\{x_n(t)\}$ von Funktionen aus $C[a, b]$ konvergiere schwach gegen die Funktion $x(t)$. Dann ist die Folge $\{x_n(t)\}$ in der Norm von $C[a, b]$ beschränkt. Zu den auf $C[a, b]$ definierten Funktionalen gehören insbesondere die Funktionale δ_{t_0}, die durch die Gleichung $\delta_{t_0}(x) = x(t_0)$ definiert waren (siehe 4.1.2., Beispiel 4). Für jedes solche Funktional bedeutet die Bedingung

$$\delta_{t_0}(x_n) \to \delta_{t_0}(x)$$

gerade

$$x_n(t_0) \to x(t_0).$$

Ist also eine Folge $\{x_n(t)\}$ schwach konvergent, so besitzt sie folgende Eigenschaften:

a) Sie ist gleichmäßig beschränkt, d. h., es gilt $|x_n(t)| \leqq c$ für alle $n = 1, 2, \ldots$ und $a \leqq t \leqq b$.

b) Sie konvergiert in jedem Punkt.

Man kann zeigen, daß diese beiden Bedingungen nicht nur notwendig, sondern auch hinreichend für die schwache Konvergenz einer Folge $\{x_n(t)\}$ in $C[a, b]$ sind. Anders ausgedrückt, stimmt die schwache Konvergenz in $C[a, b]$ (bei beschränkten Folgen) mit der punktweisen Konvergenz überein.

Offenbar stimmt diese Konvergenz nicht mit der Normkonvergenz in $C[a, b]$, d. h. mit der gleichmäßigen Konvergenz stetiger Funktionen, überein. (Man gebe ein entsprechendes Beispiel an.)

4.3.3. Die schwache Topologie und die schwache Konvergenz im dualen Raum. In 4.2.2. haben wir im dualen Raum die sogenannte starke Topologie eingeführt. Dazu hatten wir die Gesamtheit aller Mengen der Form

$$U_{\varepsilon,A} = \{f: |f(x)| < \varepsilon,\ x \in A\},$$

wobei A eine beliebige beschränkte Menge in E und ε eine beliebige positive Zahl waren, als Basis von Nullumgebungen genommen. Wenn wir hier anstelle von beschränkten Mengen nur alle endlichen Teilmengen $A \subset E$ betrachten, erhalten wir die sogenannte *schwache Topologie im dualen Raum* E^*. Da jede endliche Menge $A \subset E$ auch beschränkt ist (während die Umkehrung allgemein nicht richtig ist), ist die schwache Topologie des Raumes E^* schwächer als die starke Topologie. Im allgemeinen stimmen diese beiden Topologien nicht überein.

Die in E^* eingeführte schwache Topologie definiert in diesem Raum eine Konvergenz, die als *schwache Konvergenz von Funktionalen* bezeichnet wird. Die schwache Konvergenz von linearen Funktionalen stellt einen wichtigen Begriff dar, der in vielen Fragen der Funktionalanalysis eine wesentliche Rolle spielt. Insbesondere ist das in der Theorie der sogenannten Distributionen der Fall, die in 4.4. behandelt werden.

Die schwache Konvergenz einer Folge $\{\varphi_n\}$ von linearen Funktionalen ist offenbar gleichbedeutend mit der Konvergenz der Folge in jedem fixierten Punkt aus E. Mit anderen Worten heißt eine Folge $\{\varphi_n\}$ *schwach konvergent* gegen $\varphi \in E^*$, wenn für jedes $x \in E$ die Beziehung

$$\varphi_n(x) \to \varphi(x)$$

erfüllt ist. Trivialerweise ist auch im dualen Raum jede stark konvergente Folge schwach konvergent (aber nicht umgekehrt).

Es sei E (und folglich auch E^*) ein Banachraum. Dann gilt der folgende zu Satz 1 analoge Satz.

Satz 1*. *Ist $\{f_n\}$ eine schwach konvergente Folge linearer Funktionale auf einem Banachraum, so existiert eine Konstante C mit*

$$\|f_n\| \leqq C, \quad n = 1, 2, \ldots$$

Anders ausgedrückt: *Jede schwach konvergente Folge von Elementen aus dem dualen Raum eines Banachraumes ist bezüglich der Norm beschränkt.*

Der Beweis unterscheidet sich nicht vom Beweis von Satz 1.

Der folgende Satz ist vollkommen analog zu Satz 2.

Satz 2*. *Eine Folge linearer Funktionale* $\{\varphi_n\}$ *aus* E^* *konvergiert schwach gegen* $\varphi \in E^*$, *wenn*

1. *diese Folge beschränkt ist, d. h.*

$$\|\varphi_n\| \leqq C, \qquad n = 1, 2, \ldots;$$

2. *die Beziehung* $(\varphi_n, x) \to (\varphi, x)$ *für alle Elemente* x *einer Menge erfüllt ist, deren lineare Hülle in* E *überall dicht liegt.*

Der Beweis läßt sich genauso führen wie der Beweis von Satz 2.

Wir wollen nun ein Beispiel betrachten. E sei der Raum $C[a, b]$ der stetigen Funktionen[1]) und φ die δ-Funktion (siehe 4.1.2., Beispiel 4), d. h.

$$\varphi(x) = x(0).$$

Ferner sei $\{\varphi_n(t)\}$ eine Folge von stetigen Funktionen, die den folgenden Bedingungen genügt:

1. $\varphi_n(t) = 0$ für $|t| > \dfrac{1}{n}$, $\varphi_n(t) \geqq 0$;

2. $\int\limits_a^b \varphi_n(t)\, dt = 1$.

Dann ergibt sich für jede stetige Funktion $x(t)$ auf $[a, b]$ mit Hilfe des Mittelwertsatzes

$$\int\limits_a^b \varphi_n(t)\, x(t)\, dt = \int\limits_{-1/n}^{1/n} \varphi_n(t)\, x(t)\, dt \to x(0) \quad \text{für} \quad n \to \infty.$$

Der Ausdruck

$$\int\limits_a^b \varphi_n(t)\, x(t)\, dt$$

stellt dabei ein lineares Funktional auf $C[a, b]$ dar. Somit kann man die δ-Funktion im Sinne der schwachen Konvergenz linearer Funktionale auf $C[a, b]$ als Grenzwert einer Folge „gewöhnlicher" Funktionen darstellen.

Bemerkung. Den Raum E^* der linearen Funktionale auf einem Raum E können wir auf zwei Arten betrachten, entweder als dualen Raum des ursprünglichen Raumes E oder als Grundraum des zu E^* dualen Raumes E^{**}. Damit können wir in E^* auf zwei Arten eine schwache Topologie einführen. Einmal können wir in E^* als dem zu E dualen Raum mit Hilfe aller endlichen Systeme von Elementen aus E Umgebungen definieren. Zum anderen können wir in E^* als dem Grundraum mit Hilfe des Raumes E^{**} Umgebungen erklären. Im Fall eines reflexiven Raumes liefert das, wie nicht

[1]) Wir nehmen dabei an, daß $0 \in [a, b]$ ist. Man kann natürlich anstelle des Punktes $t = 0$ auch jeden anderen Punkt nehmen.

anders zu erwarten, genau dasselbe. Wenn E jedoch nicht reflexiv ist, erhält man auf diese Weise zwei verschiedene Topologien in E^*. Um möglichen Mißverständnissen aus dem Wege zu gehen, wollen wir die schwache Topologie im Raum der Funktionale E^* (d. h. die in E^* mit Hilfe von E definierte Topologie) als *w^*-Topologie* bezeichnen, während wir die schwache Topologie im Grundraum E^* (d. h. die mit Hilfe von E^{**} definierte Topologie) weiterhin einfach *schwache Topologie* oder *w-Topologie* nennen werden. Offensichtlich ist die w^*-Topologie in E^* schwächer als die w-Topologie des Raumes E^* (d. h., in der schwachen Topologie gibt es nicht weniger offene Mengen als in der w^*-Topologie).

4.3.4. Beschränkte Mengen im dualen Raum. In vielen Anwendungen des Begriffes der schwachen Konvergenz von linearen Funktionalen spielt der folgende Satz eine wichtige Rolle.

Satz 3. *Ist E ein separabler normierter linearer Raum, so enthält jede beschränkte Folge von stetigen linearen Funktionalen auf E eine in der w^*-Topologie konvergente Teilfolge.*

Beweis. Wir wählen in E eine überall dichte Menge $(x_1, x_2, \ldots, x_n, \ldots)$. Ist $\{\varphi_n\}$ eine (in der Norm) beschränkte Folge linearer Funktionale auf E, so ist die Zahlenfolge

$$\varphi_1(x_1), \varphi_2(x_1), \ldots, \varphi_n(x_1), \ldots$$

beschränkt. Daher kann man eine Teilfolge

$$\varphi_1^{(1)}, \varphi_2^{(1)}, \ldots, \varphi_n^{(1)}, \ldots$$

auswählen, so daß die Zahlenfolge

$$\varphi_1^{(1)}(x_1), \varphi_2^{(1)}(x_1), \ldots, \varphi_n^{(1)}(x_1), \ldots$$

konvergent ist. Ferner kann man aus der Folge $\{\varphi_n^{(1)}\}$ eine Teilfolge

$$\varphi_1^{(2)}, \varphi_2^{(2)}, \ldots, \varphi_n^{(2)}, \ldots$$

auswählen, so daß die Zahlenfolge

$$\varphi_1^{(2)}(x_2), \varphi_2^{(2)}(x_2), \ldots, \varphi_n^{(2)}(x_2), \ldots$$

konvergent ist. Wenn wir diesen Prozeß fortsetzen, erhalten wir ein System

$$\varphi_1^{(1)}, \varphi_2^{(1)}, \ldots, \varphi_n^{(1)}, \ldots,$$

$$\varphi_1^{(2)}, \varphi_2^{(2)}, \ldots, \varphi_n^{(2)}, \ldots,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots$$

von Folgen (von denen jede eine Teilfolge der vorhergehenden Folge ist), wobei $\{\varphi_n^{(k)}\}$ in den Punkten $x_1, x_2, \ldots, x_k$ konvergent ist. Nehmen wir nun die Diagonalfolge

$$\varphi_1^{(1)}, \varphi_2^{(2)}, \ldots, \varphi_n^{(n)}, \ldots,$$

so erhalten wir eine Teilfolge der ursprünglichen Folge $\{\varphi_n\}$, für die die Zahlenfolge

$$\varphi_1{}^{(1)}(x_n),\ \varphi_2{}^{(2)}(x_n),\ \dots$$

in jedem Punkt x_n konvergent ist. Dann konvergiert aber die Folge $\varphi_1{}^{(1)}(x)$, $\varphi_2{}^{(2)}(x)$, ... (nach Satz 2*) auch für jedes $x \in E$. Damit ist der Satz bewiesen.

Dieser Satz besagt zusammen mit Satz 1*, daß im dualen Raum E^* eines separablen Banachraumes E die beschränkten Mengen und nur diese in der w^*-Topologie abzählbar-präkompakt sind. Wir werden zeigen, daß die beschränkten Mengen hier sogar präkompakt und nicht nur abzählbar-präkompakt sind. Dazu beweisen wir zunächst folgenden Satz.

Satz 4. *E sei ein separabler normierter Raum, und S^* sei die Einheitskugel des zu E dualen Raumes E^*. Dann kann die in S^* durch die w^*-Topologie von E^* erzeugte Topologie mit Hilfe der Metrik*

$$\varrho(f, g) = \sum 2^{-n} \,|(f - g, x_n)| \tag{2}$$

erklärt werden; dabei ist $\{x_n\}$ eine fixierte abzählbare überall dichte Menge in der Einheitskugel S des Raumes E.

Beweis. Es ist klar, daß die Funktion $\varrho(f, g)$ alle Eigenschaften eines Abstandes besitzt. Außerdem ist sie translationsinvariant,

$$\varrho(f + h, g + h) = \varrho(f, g).$$

Deshalb brauchen wir nur die Äquivalenz der beiden Nullumgebungssysteme nachzuweisen, die in S^* einmal durch die w^*-Topologie des Raumes E^* und zum anderen durch die Metrik (2) definiert werden. Dazu müssen wir zeigen, daß a) jede Kugel

$$Q_\varepsilon = \{f : \varrho(f, 0) < \varepsilon\}$$

den Durchschnitt von S^* mit einer w^*-Nullumgebung enthält und daß b) jede w^*-Nullumgebung in E^* den Durchschnitt von S^* mit einer Kugel Q_ε enthält.

Zunächst wollen wir a) zeigen. Dazu wählen wir ein N mit $2^{-N} < \varepsilon/2$ und betrachten die w^*-Nullumgebung

$$V = V_{x_1,\dots,x_N;\varepsilon/2} = \left\{f : |(f, x_k)| < \frac{\varepsilon}{2},\ k = 1, 2, \dots, N\right\}.$$

Für jedes $f \in S^* \cap V$ folgt dann

$$\varrho(f, 0) = \sum_{n=1}^{N} 2^{-n} \,|(f, x_n)| + \sum_{n=N+1}^{\infty} 2^{-n} \,|(f, x_n)|$$

$$\leqq \frac{\varepsilon}{2} \sum_{n=1}^{N} 2^{-n} + \sum_{n=N+1}^{\infty} 2^{-n} < \varepsilon,$$

d. h. $S^* \cap V \subset Q_\varepsilon$. Somit ist die Behauptung a) bewiesen. Zum Beweis von b) sei

$$U = U_{y_1,\ldots,y_m;\delta} = \{f\colon |(f, y_k)| < \delta,\ k = 1, 2, \ldots, m\}$$

eine w^*-Nullumgebung in E^*. Dabei kann man dann $\|y_k\| \leqq 1$, $k = 1, 2, \ldots, m$, annehmen. Da die Menge $\{x_n\}$ überall in S dicht ist, gibt es Indizes $n_1, \ldots, n_m$ mit $\|y_k - x_{n_k}\| < \delta/2$ für $k = 1, 2, \ldots, m$. Wir setzen nun $N = \max(n_1, \ldots, n_m)$ und $\varepsilon = 2^{-(N+1)}\delta$. Dann ergibt sich für jedes $f \in S^* \cap Q_\varepsilon$ aus der Ungleichung

$$\sum_{n=1}^{\infty} 2^{-n} |(f, x_n)| < \varepsilon$$

die Beziehung $|(f, x_n)| < 2^n\varepsilon$, insbesondere also

$$|(f, x_{n_k})| < 2^{n_k}\varepsilon \leqq 2^N\varepsilon = \frac{\delta}{2}.$$

Folglich erhalten wir für alle $k = 1, 2, \ldots, m$

$$\begin{aligned} |(f, y_k)| &\leqq |(f, x_{n_k})| + |(f, y_k - x_{n_k})| \\ &< \frac{\delta}{2} + \|f\| \cdot \|y_k - x_{n_k}\| < \delta. \end{aligned}$$

Somit gilt $A^* \cap Q_\varepsilon \subset U$, was zu zeigen war.

Es ist klar, daß sich dieses Resultat automatisch auf eine beliebige Kugel und damit auf jede beschränkte Teilmenge $M \subset E^*$ überträgt.

Wir haben gezeigt (Satz 3), daß man aus jeder beschränkten Folge in E^* eine w^*-konvergente Teilfolge auswählen kann. Anders ausgedrückt: Jede beschränkte Teilmenge M im dualen Raum E^* eines separablen normierten linearen Raumes ist abzählbar-präkompakt bezüglich der w^*-Topologie. Auf Grund des letzten Satzes bildet aber jede solche Menge einen metrisierbaren topologischen Raum, und für metrische Räume stimmen Kompaktheit und abzählbare Kompaktheit überein. Somit erhalten wir das folgende Resultat.

Satz 3*. *Jede beschränkte Menge M im dualen Raum E^* eines separablen normierten Raumes ist präkompakt in der w^*-Topologie.*

Wir werden jetzt zeigen, daß jede abgeschlossene Kugel des Raumes (E^*, b) in der w^*-Topologie abgeschlossen ist, falls E ein separabler normierter linearer Raum ist. Jede Translation im Raum E^* führt die Klasse der (bezüglich der w^*-Topologie) abgeschlossenen Mengen in sich über. Also brauchen wir nur zu beweisen, daß jede Kugel der Form $S_c^* = \{f\colon \|f\| \leqq c\}$ in der w^*-Topologie abgeschlossen ist.

Dazu sei $f_0 \notin S_c^*$. Nach Definition der Norm eines Funktionals gibt es einen Punkt $x \in E$ mit $\|x\| = 1$ und $f_0(x) = \alpha > c$. Dann bildet die Menge $U = \left\{f\colon f(x) > \dfrac{\alpha + c}{2}\right\}$ eine w^*-Umgebung des Funktionals f_0, die kein Element der Kugel S_c^* enthält. Also ist die Kugel S_c^* in der w^*-Topologie abgeschlossen.

Aus der bewiesenen Behauptung und aus Satz 3* ergibt sich der folgende Satz.

Satz 5. *Jede abgeschlossene Kugel im dualen Raum eines separablen normierten Raumes ist kompakt in der w^*-Topologie.*

Die oben angeführten Resultate über beschränkte Mengen in dualen Räumen können von normierten Räumen auf beliebige lokalkonvexe Räume übertragen werden. Siehe dazu etwa [52].

4.4. Distributionen

4.4.1. Erweiterung des Funktionsbegriffs. In den verschiedenen Problemen der Analysis wird der Terminus „Funktion" verschieden allgemein aufgefaßt. Manchmal werden stetige Funktionen betrachtet, in anderen Fragen muß man voraussetzen, daß man es mit ein- oder mehrmals differenzierbaren Funktionen zu tun hat, usw. Jedoch in einer Reihe von Fällen erweist sich der klassische Funktionsbegriff als unzureichend, welbst wenn er im weitesten Sinne verstanden wird, d. h. als Vorschrift, die jedem Wert x aus dem Definitionsbereich dieser Funktion eine gewisse Zahl $y = f(x)$ zuordnet. Dazu geben wir zwei wichtige Beispiele an.

1. Eine Massenverteilung auf der Zahlengeraden kann auf bequeme Weise mittels der Dichte dieser Verteilung angegeben werden. Wenn jedoch Punkte auf der Zahlengeraden existieren, die eine positive Masse tragen, kann die Dichte dieser Verteilung offenbar nicht durch eine „gewöhnliche" Funktion beschrieben werden.

2. Bei Anwendung der Analysis auf gewisse Probleme stellt man fest, daß manche Operationen nicht immer ausführbar sind. zum Beispiel kann man eine Funktion, die (in einigen Punkten oder überall) keine Ableitung besitzt, nicht (im üblichen Sinne) differenzieren. Natürlich kann man Schwierigkeiten dieser Art aus dem Wege gehen, indem man sich etwa nur auf die Betrachtung von analytischen Funktionen beschränkt. Aber eine solche Einschränkung des zugelassenen Funktionenbereichs ist in vielen Fällen nicht sehr wünschenswert.

Es zeigt sich jedoch, daß man derartige Schwierigkeiten überwinden kann, indem man den Funktionsbegriff nicht einengt, sondern wesentlich erweitert und die sogenannten Distributionen einführt. Als Grundlage für die Einführung der entsprechenden Definitionen dient der oben betrachtete Begriff des dualen Raumes.

Wir betonen noch einmal, daß die Einführung von Distributionen keineswegs durch das Streben nach möglichst weiter Ausdehnung der Begriffe der Analysis, sondern durch konkrete Aufgaben hervorgerufen wurde. Im Prinzip wurden in der Physik schon ziemlich lange Distributionen benutzt, jedenfalls noch bevor die streng mathematische Theorie der Distributionen entstand.

Ehe wir zu den exakten Definitionen übergehen, wollen wir den Grundgedanken erläutern, der zu den Distributionen führt.

Es sei f eine feste lokalintegrierbare Funktion auf der reellen Achse, d. h., f sei auf jedem beschränkten Intervall integrierbar. Außerdem sei φ eine stetige Funktion,

die außerhalb eines gewissen endlichen Intervalls verschwindet (solche Funktionen werden wir im weiteren *finit* nennen). Jeder solchen Funktion φ kann man mit Hilfe der fixierten Funktion f die Zahl

$$(f, \varphi) = \int_{-\infty}^{\infty} f(x)\, \varphi(x)\, dx \tag{1}$$

zuordnen (hier wird wegen der Finitheit von $\varphi(x)$ tatsächlich nur über ein endliches Intervall integriert). Anders ausgedrückt, kann man die Funktion f als Funktional (und zwar als lineares, wegen der Grundeigenschaften des Integrals) auf einem gewissen Raum von finiten Funktionen betrachten. Jedoch stellen die Funktionale der Gestalt (1) nicht alle Funktionale dar, die man auf einem solchen Raum einführen kann. Wenn wir zum Beispiel jeder Funktion φ ihren Wert im Punkt 0 zuordnen, erhalten wir ein lineares Funktional, das nicht in der Form (1) darstellbar ist. Somit werden die Funktionen $f(x)$ auf natürliche Weise in eine größere Menge eingebettet, und zwar in die Gesamtheit aller linearen Funktionale auf den finiten Funktionen.

Den Bereich der Funktionen φ kann man verschieden wählen. Zum Beispiel könnte man alle stetigen finiten Funktionen nehmen. Wie jedoch im weiteren klar wird, ist es vernünftig, von den zulässigen Funktionen φ neben der Stetigkeit und der Finitheit auch noch hinreichend starke Glattheitsbedingungen zu verlangen.

4.4.2. Der Raum der Grundfunktionen. Wir kommen jetzt zu den exakten Definitionen. Wir betrachten die Gesamtheit K aller finiten, unendlich oft differenzierbaren Funktionen auf der reellen Achse.[1]) Die zu K gehörenden Funktionen bilden einen linearen Raum (mit der üblichen Addition und Skalarmultiplikation von Funktionen). In diesem Raum kann man zwar keine Norm im Sinne der oben entwickelten Theorie einführen, jedoch läßt sich in natürlicher Weise ein Konvergenzbegriff erklären.

Eine Folge $\{\varphi_n\}$ von Elementen aus K heißt *konvergent gegen die Funktion* $\varphi \in K$, wenn a) ein Intervall existiert, auf dessen Komplement alle φ_n identisch verschwinden[1]) und wenn b) die Folge der k-ten Ableitungen[2]) $\{\varphi_n^{(k)}\}$ $(k = 0, 1, 2, \ldots)$ auf diesem Intervall gleichmäßig gegen $\varphi^{(k)}$ konvergiert.

Den linearen Raum K mit der oben definierten Konvergenz werden wir *Grundraum* und seine Elemente *Grundfunktionen* oder *Testfunktionen* nennen.

Es ist nicht schwer, die Topologie in K zu beschreiben, die der in K erklärten Konvergenz entspricht. Diese Topologie wird durch das folgende Nullumgebungssystem erzeugt. Eine Nullumgebung besteht aus allen Funktionen von K, die für alle x den Ungleichungen

$$|\varphi(x)| < \gamma_0(x), \ldots, |\varphi^{(m)}(x)| < \gamma_m(x)$$

[1]) Dabei kann das Intervall, auf dessen Komplement die Funktion φ verschwindet, für verschiedene $\varphi \in K$ verschieden sein.

[2]) Unter der nullten Ableitung versteht man üblicherweise die Funktion selbst.

genügen, wobei $\gamma_0, \ldots, \gamma_m$ endlich viele positive stetige Funktionen sind. Der Nachweis, daß diese Topologie tatsächlich der oben beschriebenen Konvergenz entspricht, wird dem Leser überlassen.

Aufgabe. Mit K_m bezeichnen wir den Teilraum von K, der aus allen Funktionen $\varphi \in K$ mit $\varphi(x) = 0$ für $|x| > m$ besteht. Der Raum K_m wird zu einem abzählbar-normierten Raum, wenn wir die Normen

$$\|\varphi\|_n = \sup_{\substack{0 \leq k \leq n \\ |x| \leq m}} |\varphi^{(k)}(x)|, \qquad n = 0, 1, 2, \ldots,$$

einführen. Man beweise die folgende Aussage: Die in K_m durch diese Normen erzeugte Topologie (bzw. Folgenkonvergenz) stimmt mit der Topologie (bzw. Konvergenz) überein, die in K_m durch die oben beschriebene Topologie (Konvergenz) des Raumes K induziert wird. Offenbar ist $K_1 \subset K_2 \subset \cdots \subset K_m \subset$ und $K = \bigcup_{m=1}^{\infty} K_m$. Man zeige, daß eine Menge $Q \subset K$ dann und nur dann beschränkt bezüglich der in K eingeführten Topologie ist, wenn ein m existiert, so daß Q eine beschränkte Teilmenge des abzählbar-normierten Raumes K_m ist. Schließlich sei T ein lineares Funktional auf K, und man beweise, daß die folgenden vier Aussagen gleichwertig sind:

a) Das Funktional T ist bezüglich der Topologie des Raumes K stetig.

b) Das Funktional T ist auf jeder beschränkten Menge $Q \subset K$ beschränkt.

c) Aus $\varphi_n \in K$ und $\varphi_n \to 0$ (im Sinne der in K erklärten Konvergenz von Folgen) folgt $T(\varphi_n) \to 0$.

d) Für jedes m ist die Einschränkung T_m des Funktionals T auf den Teilraum $K_m \subset K$ ein stetiges Funktional auf K_m.

4.4.3. Distributionen

Definition 1. Jedes stetige lineare Funktional $T(\varphi)$ auf dem Grundraum K heißt *Distribution* (auf der reellen Achse $-\infty < x < \infty$). Die Stetigkeit ist dabei so zu verstehen, daß $T(\varphi_n) \to T(\varphi)$ strebt, wenn die Folge φ_n im Grundraum K gegen φ konvergiert.

Wir vermerken zunächst erst einmal, daß jede lokalintegrierbare Funktion durch die Gleichung

$$T_f(\varphi) = \int_{-\infty}^{\infty} f(x)\, \varphi(x)\, dx \tag{2}$$

ein stetiges lineares Funktional auf K, d. h. eine Distribution definiert.[1]) Derartige Distributionen werden wir im weiteren *regulär* nennen, während wir alle verbleiben-

[1]) Anm. d. Übers.: Sind f_1 und f_2 zwei lokalintegrierbare Funktionen, so gilt

$$T_{f_1}(\varphi) = T_{f_2}(\varphi) \quad \text{für alle } \varphi \in K$$

genau dann, wenn die Menge $\{x: f_1(x) \neq f_2(x)\}$ das Lebesguesche Maß Null hat. Anders ausgedrückt: $T_{f_1} = T_{f_2}$ gilt genau dann, wenn f_1 äquivalent zu f_2 ist. Identifiziert man Funktionen, die sich nur auf einer Menge vom Maß Null unterscheiden, so erhält man eine eineindeutige Zuordnung zwischen den Distributionen des obigen Typs (den regulären Distributionen) und den lokalintegrierbaren Funktionen. Hiervon wird später Gebrauch gemacht. Man vergleiche hierzu die Ausführungen in 4.4.5.

den Distributionen, die also nicht in der Form (2) darstellbar sind, *singulär* nennen werden.

Wir wollen einige Beispiele singulärer Distributionen anführen.

1. *Die δ-Funktion.* Die Gleichung

$$T(\varphi) = \varphi(0)$$

definiert ein stetiges lineares Funktional auf K, also gemäß unserer oben eingeführten Terminologie eine Distribution. Dieses Funktional schreibt man gewöhnlich in der Form

$$\int\limits_{-\infty}^{\infty} \delta(x)\, \varphi(x)\, dx. \tag{3}$$

Dabei versteht man unter $\delta(x)$ eine „Funktion“, die für $x \neq 0$ verschwindet und im Punkt $x = 0$ unendlich wird, und zwar so, daß

$$\int\limits_{-\infty}^{\infty} \delta(x)\, dx = 1$$

ist. Wir haben die δ-Funktion schon in 4.1. als Funktional auf dem Raum aller stetigen Funktionen über einem Intervall betrachtet. Die Betrachtung der δ-Funktion als Funktional auf K besitzt jedoch bestimmte Vorzüge, z. B. kann man dann den Begriff der Ableitung für die δ-Funktion erklären.

2. *Die verschobene δ-Funktion.* Es sei

$$T(\varphi) = \varphi(a).$$

Dieses Funktional schreibt man in Analogie zur Bezeichnungsweise (3) auch in der Form

$$\int\limits_{-\infty}^{\infty} \delta(x-a)\, \varphi(x)\, dx. \tag{4}$$

3. *Die Ableitung der δ-Funktion.* Jedem $\varphi \in K$ ordnen wir jetzt die Zahl $-\varphi'(0)$ zu. Später werden wir klären, warum man dieses Funktional natürlicherweise als Ableitung des Funktionals aus dem ersten Beispiel ansieht.

4. Wir betrachten die Funktion $1/x$. Sie ist auf keinem Intervall integrierbar, das den Nullpunkt enthält. Jedoch für jedes $\varphi \in K$ existiert der Hauptwert des Integrals

$$\int\limits_{-\infty}^{\infty} \varphi(x) \frac{1}{x}\, dx$$

und ist endlich. Es gilt nämlich

$$\int_{-\infty}^{\infty} \varphi(x) \frac{1}{x} \, dx = \int_{a}^{b} \varphi(x) \frac{1}{x} \, dx = \int_{a}^{b} \frac{\varphi(x) - \varphi(0)}{x} \, dx + \int_{a}^{b} \frac{\varphi(0)}{x} \, dx,$$

dabei ist (a, b) das Intervall, außerhalb dessen φ verschwindet. Das erste der rechts stehenden Integrale existiert hier im gewöhnlichen Sinne (unter dem Integral steht eine stetige Funktion), während das zweite Integral im Sinne des Hauptwertes existiert. Somit erzeugt $1/x$ ein Funktional auf K, d. h. eine Distribution.

Man kann beweisen, daß alle Distributionen aus den Beispielen 1 bis 4 singulär sind (d. h. sich nicht in der Form (2) mit irgendeiner lokalintegrierbaren Funktion f darstellen lassen).

4.4.4. Operationen im Bereich der Distributionen. Für Distributionen, d. h. für stetige lineare Funktionale auf K, sind eine Addition und eine Skalarmultiplikation erklärt. Im Fall regulärer Distributionen (die ja auch als gewöhnliche Funktionen aufgefaßt werden können) stimmt diese Addition mit der gewöhnlichen Addition von Funktionen überein. Genauso verhält es sich auch mit der Skalarmultiplikation.

Wir wollen im Raum der Distributionen einen Konvergenzbegriff einführen, und zwar nennen wir eine Folge $\{\varphi_n\}$ von Distributionen *konvergent gegen* f, wenn für jedes $\varphi \in K$ die Beziehung

$$(f_n, \varphi) \to (f, \varphi)$$

erfüllt ist. Anders ausgedrückt: Wir definieren die Konvergenz einer Folge von Distributionen als punktweise Konvergenz in jedem Element aus K. Den Raum der Distributionen mit dieser Konvergenz bezeichnen wir mit K^*.

Ist α eine unendlich oft differenzierbare Funktion, so läßt sich das Produkt von α mit einer Distribution f in natürlicher Weise durch die Formel

$$(\alpha f, \varphi) = (f, \alpha\varphi)$$

definieren (der rechts stehende Ausdruck besitzt wegen $\alpha\varphi \in K$ einen Sinn). Alle diese Operationen, die Addition, die Skalarmultiplikation und die Multiplikation mit einer unendlich oft differenzierbaren Funktion, sind stetig.

Ein Produkt zweier Distributionen führen wir nicht ein. Sinnvollerweise müßte diese Operation nämlich stetig sein und für zwei reguläre Distributionen mit der üblichen Multiplikation von Funktionen übereinstimmen. Man kann aber zeigen, daß es ein solches Produkt nicht geben kann.

Wir werden jetzt für Distributionen eine Differentiation erklären und deren Eigenschaften untersuchen.

Zunächst sei T eine reguläre Distribution, die durch eine stetig differenzierbare Funktion f erzeugt wird,

$$T(\varphi) = \int_{-\infty}^{\infty} f(x)\, \varphi(x)\, dx.$$

Als deren Ableitung möchte man natürlich das durch die Formel

$$\frac{dT}{dx}(\varphi) = \int_{-\infty}^{\infty} f'(x)\, \varphi(x)\, dx$$

definierte Funktional $\frac{dT}{dx}$ erhalten. Durch partielle Integration ergibt sich daraus

$$\frac{dT}{dx}(\varphi) = \int_{-\infty}^{\infty} f'(x)\, \varphi(x)\, dx = -\int_{-\infty}^{\infty} f(x)\, \varphi'(x)\, dx.$$

Dabei haben wir benutzt, daß φ außerhalb eines gewissen endlichen Intervalls verschwindet. Somit haben wir für $\frac{dT}{dx}$ einen Ausdruck erhalten, in dem die Ableitung der Funktion f nicht vorkommt. Diese Überlegungen führen zu der folgenden Definition:

Definition 2. Ist T eine Distribution, so heißt das durch die Formel

$$\frac{dT}{dx}(\varphi) = -T(\varphi')$$

definierte Funktional *Ableitung* von T. Es ist klar, daß das so definierte Funktional linear und stetig ist, d. h. eine Distribution darstellt. Analog definiert man die zweite, dritte und noch höhere Ableitungen.

Wenn wir die Distribution mit dem Symbol f bezeichnen, so werden wir die Ableitung f (die wir in dem gerade erklärten Sinne verstehen) mit dem üblichen Symbol f' bezeichnen.

Aus der Definition der Ableitung einer Distribution ergibt sich unmittelbar die Gültigkeit der folgenden Aussagen:

1. *Jede Distribution besitzt Ableitungen beliebig hoher Ordnung.*

2. *Konvergiert die Folge* $\{f_n\}$ *von Distributionen gegen die Distribution* f (im Sinne der Konvergenz von Distributionen), *dann konvergiert die Folge* $\{f_n'\}$ *der Ableitungen gegen die Ableitung* f' *von* f. *Das gleiche gilt auch für die Ableitungen beliebiger Ordnung.* Das ist gleichwertig damit, daß *man jede konvergente Reihe von Distributionen beliebig oft gliedweise differenzieren kann.*

Wir wollen einige Beispiele betrachten.

1. Ist f eine differenzierbare Funktion mit stetiger (oder stückweise stetiger) Ableitung, so ist nach dem oben Gesagten klar, daß die Ableitung von f im Sinne der Distributionen mit der gewöhnlichen Ableitung übereinstimmt.[1])

2. Es sei

$$f(x) = \begin{cases} 1 \text{ für } x > 0, \\ 0 \text{ für } x \leqq 0. \end{cases} \tag{5}$$

Diese Funktion wird *Heaviside-Funktion* genannt, sie definiert das lineare Funktional

$$(f, \varphi) = \int_0^\infty \varphi(x)\, dx.$$

Nach Definition der Distributionsableitung ergibt sich daraus

$$(f', \varphi) = -(f, \varphi') = -\int_0^\infty \varphi'(x)\, dx = \varphi(0)$$

(da φ im Unendlichen verschwindet). Somit erhält man als Ableitung der Heaviside-Funktion (5) gerade die δ-Funktion.

3. Ist f eine Funktion, die in den Punkten $x_1, x_2, \ldots$ Sprünge der Höhe $h_1, h_2, \ldots$ besitzt und in allen anderen Punkten differenzierbar ist, so ergibt sich die Distributionsableitung von f nach den Beispielen 1 und 2 offenbar als Summe von f' (in den Punkten, in denen diese Ableitung existiert) und eines Ausdrucks der Form

$$\sum_i h_i \delta(x - x_i).$$

4. Die Ableitung der δ-Funktion im Sinne der Distributionen ist das Funktional, das jeder Funktion φ aus K den Wert $-\varphi'(0)$ zuordnet. Das ist aber gerade jenes Funktional, das wir schon „Ableitung der δ-Funktion" genannt hatten.

5. Wir betrachten die Reihe

$$\sum_{n=1}^\infty \frac{\sin nx}{n}. \tag{6}$$

Ihre Summe ist eine Funktion mit der Periode 2π, die im Intervall $[-\pi, \pi]$ die Gestalt

$$f(x) = \begin{cases} \dfrac{\pi - x}{2} & \text{für} \quad 0 < x \leqq \pi, \\ -\dfrac{\pi + x}{2} & \text{für} \quad -\pi \leqq x < 0, \\ 0 & \text{für} \quad x = 0 \end{cases}$$

[1]) Anm. d. Übers.: Man vergleiche die Fußnote auf S. 205.

besitzt. Die verallgemeinerte Ableitung dieser Funktion ergibt die Distribution

$$-\frac{1}{2} + \pi \sum_{k=-\infty}^{\infty} \delta(x - 2k\pi) \tag{7}$$

(wenn wir sie auf eine beliebige finite Funktion anwenden, erhalten wir immer nur endlich viele von Null verschiedene Summanden). Andererseits erhalten wir durch gliedweise Differentiation der Reihe (6) die divergente Reihe

$$\sum_{n=1}^{\infty} \cos nx.$$

Im Sinne der Konvergenz von Distributionen ist diese Reihe jedoch konvergent (und zwar gegen den Ausdruck (7)). Mit Hilfe der Distributionen ist es also möglich, der Summe einer im gewöhnlichen Sinne divergenten Reihe einen ganz bestimmten Sinn zu geben. Genauso verhält es sich auch mit vielen divergenten Integralen. Diesem Umstand begegnet man oft in der Quantenfeldtheorie und in einigen anderen Gebieten der theoretischen Physik. Im übrigen entsteht eine ähnliche Situation schon bei der Lösung von elementaren Aufgaben der mathematischen Physik mit Hilfe der Fourierschen Methode. So tauchen etwa bei der Lösung der Gleichung $\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}$ (die zum Beispiel die Schwingungen einer Saite beschreibt) trigonometrische Reihen auf, die nur im Sinne der Theorie der Distributionen zweite Ableitungen nach x und t besitzen und somit dieser Gleichung nur im Sinne der Distributionen genügen.

4.4.5. Bemerkungen zum Raum der Grundfunktionen. Wir haben die Distributionen als lineare Funktionale auf einem gewissen Raum definiert, und zwar auf dem Raum K der finiten unendlich oft differenzierbaren Funktionen. Man hätte den Grundraum aber auch anders wählen können. Wir wollen die Gründe untersuchen, warum gerade der Raum K als Raum der Grundfunktionen genommen wurde. Derartige Überlegungen kann man auch in den anderen Fällen anstellen.

Durch die scharfen Forderungen, daß jedes Element aus K finit und beliebig oft differenzierbar sein sollte, haben wir erstens einen umfangreichen Raum von Distributionen erhalten (eine Einschränkung des Grundraumes führt offenbar zur Vergrößerung des dualen Raumes). Und zweitens konnten auf diese Weise die Grundoperationen der Analysis (Grenzübergang, Differentiation) auf Distributionen übertragen werden. Dabei ist der Raum K der Grundfunktionen aber auch nicht zu eng. In K gibt es hinreichend viele Elemente, um etwa mit ihrer Hilfe die stetigen Funktionen zu trennen. Darunter ist folgendes zu verstehen. *Es seien f_1 und f_2 zwei verschiedene stetige* (und somit auch lokalintegrierbare) *Funktionen auf der reellen Achse. Dann*

existiert eine Funktion $\varphi \in K$ mit

$$\int_{-\infty}^{\infty} f_1(x)\, \varphi(x)\, dx \neq \int_{-\infty}^{\infty} f_2(x)\, \varphi(x)\, dx. \tag{8}$$

Um das zu zeigen, setzen wir $f(x) = f_1(x) - f_2(x)$. Ist $f(x) \not\equiv 0$, so existiert ein Punkt x_0 mit $f(x_0) \neq 0$. Dann ändert aber $f(x)$ in einem ganzen Intervall (α, β), das den Punkt x_0 enthält, sein Vorzeichen nicht. Nun betrachten wir die Funktion

$$\varphi(x) = \begin{cases} e^{-\frac{1}{(\beta - x)(x - \alpha)}} & \text{für } \alpha < x < \beta, \\ 0 & \text{sonst.} \end{cases}$$

Diese Funktion gehört zu K, denn sie verschwindet außerhalb des Intervalls (α, β), und sie besitzt Ableitungen jeder Ordnung. (Man zeige die Existenz der Ableitungen in den Punkten $x = \alpha$ und $x = \beta$!) Außerdem ist $\varphi(x)$ im Intervall (α, β) positiv, woraus offenbar

$$\int_{-\infty}^{\infty} f(x)\, \varphi(x)\, dx = \int_{\alpha}^{\beta} f(x)\, \varphi(x)\, dx \neq 0$$

folgt. Somit haben wir gezeigt, daß der Raum K ausreicht, um je zwei stetige Funktionen zu trennen.[1])

4.4.6. Die Bestimmung von Distributionen aus ihren Ableitungen. Differentialgleichungen in der Klasse der Distributionen. Differentialgleichungen sind eines der Hauptanwendungsgebiete der Theorie der Distributionen. Zugleich haben Aufgaben, die mit solchen Gleichungen zusammenhängen, auch in bedeutendem Maße die Entwicklung dieser Theorie stimuliert. Hauptsächlich wird diese Theorie auf partielle Differentialgleichungen angewendet, die wir aber hier nicht betrachten wollen. Wir werden hier nur einfache Fragen berühren, die sich auf die Lösung gewöhnlicher Differentialgleichungen mit Hilfe von Distributionen beziehen.

Wir beginnen mit der einfachsten Gleichung dieser Form,

$$y' = f(x)$$

($f(x)$ ist hier eine gewöhnliche Funktion oder eine Distribution), d. h., wir wollen eine Distribution aus ihrer Ableitung bestimmen. Zunächst sei $f(x) \equiv 0$.

---

[1]) Diese Behauptung kann man auf wesentlich allgemeinere Funktionen als die stetigen Funktionen ausdehnen. Dazu benötigen wir aber den Begriff der Lebesgue-Integrierbarkeit, von dem erst im folgenden Kapitel die Rede sein wird. [Anm. d. Übers.: Man vergleiche hierzu die Fußnote auf S. 205.]

Satz 1. *Nur die Konstanten sind Lösungen der Gleichung*

$$y' = 0 \tag{9}$$

(*in der Klasse der Distributionen*).

Beweis. Gleichung (9) bedeutet, daß für jede Grundfunktion $\varphi \in K$

$$(y', \varphi) = (y, -\varphi') = 0 \tag{10}$$

ist. Mit $K^{(1)}$ bezeichnen wir die Gesamtheit aller der Grundfunktionen, die als Ableitung einer Grundfunktion dargestellt werden können. Offenbar ist $K^{(1)}$ ein linearer Teilraum von K. Setzen wir $\varphi_1(x) = -\varphi'(x)$, so durchläuft φ_1 den Raum $K^{(1)}$, wenn φ den Raum K durchläuft. Gleichung (10) definiert also ein Funktional y auf $K^{(1)}$.

Eine Grundfunktion φ gehört nun genau dann zu $K^{(1)}$, wenn

$$\int_{-\infty}^{\infty} \varphi(x)\, dx = 0 \tag{11}$$

ist, d. h., $K^{(1)}$ ist der Kern des Funktionals $\int_{-\infty}^{\infty} \varphi(x)\, dx$. Denn aus $\varphi(x) = \psi'(x)$ folgt

$$\int_{-\infty}^{\infty} \varphi(x)\, dx = \psi(x) \Big|_{-\infty}^{\infty} = 0. \tag{12}$$

Umgekehrt ist

$$\psi(x) = \int_{-\infty}^{x} \varphi(t)\, dt \tag{13}$$

eine beliebig oft differenzierbare Funktion mit $\psi'(x) = \varphi(x)$, und diese Funktion ist finit, wenn (11) erfüllt ist.

Nach den Resultaten aus 3.1.6. kann man also eine beliebige Grundfunktion $\varphi \in K$ in der Form

$$\varphi = \varphi_1 + c \cdot \varphi_0 \qquad (\varphi_1 \in K^{(1)})$$

darstellen, wobei φ_0 eine feste Grundfunktion ist mit $\varphi_0 \notin K^{(1)}$ und

$$\int_{-\infty}^{\infty} \varphi_0(x)\, dx = 1.$$

Dazu braucht man nur

$$c = \int_{-\infty}^{\infty} \varphi(x)\, dx \qquad \text{und} \qquad \varphi_1(x) = \varphi(x) - c\varphi_0(x)$$

zu setzen. Somit ist das Funktional y durch die Angabe seines Wertes für die Grundfunktion $\varphi_0(x)$ bereits eindeutig auf ganz K bestimmt. Setzen wir also $(y, \varphi_0) = \alpha$,

so erhalten wir

$$(y, \varphi) = (y, \varphi_1) + c(y, \varphi_0) = \alpha \int_{-\infty}^{\infty} \varphi(x)\, dx = \int_{-\infty}^{\infty} \alpha\varphi(x)\, dx,$$

d. h., die Distribution y ist identisch mit der Konstanten α.[1]) Das war aber gerade zu beweisen.

Hieraus folgt, daß $f - g = \text{const}$ ist, wenn die beiden Distributionen f und g die Gleichung $f' = g'$ erfüllen.

Jetzt betrachten wir die Gleichung

$$y' = f(x), \tag{14}$$

in der $f(x)$ eine beliebige Distribution ist.

Satz 2. *Gleichung* (14) *besitzt für jedes $f \in K^*$ eine Lösung, die wieder zu K^* gehört*

Diese Lösung bezeichnet man natürlicherweise als *Integral* der Distribution f.

Beweis. Gleichung (14) besagt, daß für jede Grundfunktion φ

$$(y', \varphi) = (y, -\varphi') = (f, \varphi) \tag{15}$$

gilt. Diese Gleichheit bestimmt den Wert des Funktionals y auf allen Grundfunktionen φ_1 aus $K^{(1)}$:

$$(y, \varphi_1) = \left(f, -\int_{-\infty}^{x} \varphi_1(\xi)\, d\xi\right).$$

Jetzt benutzen wir die oben erhaltene Darstellung

$$\varphi = \varphi_1 + c\varphi_0$$

für Elemente aus K. Setzen wir $(y, \varphi_0) = 0$, so ist das Funktional y auf ganz K erklärt, und zwar ist

$$(y, \varphi) = (y, \varphi_1) = \left(f, -\int_{-\infty}^{x} \varphi_1(\xi)\, d\xi\right).$$

Dieses Funktional ist linear und stetig, wie man leicht überprüfen kann, und erfüllt außerdem Gleichung (14). Denn es gilt für alle $\varphi \in K$

$$(y', \varphi) = (y, -\varphi') = \left(f, \int_{-\infty}^{x} \varphi'(\xi)\, d\xi\right) = (f, \varphi).$$

Somit existiert für jede Distribution $f(x)$ eine Lösung der Gleichung

$$y' = f(x),$$

d. h., jede Distribution besitzt ein Integral. Nach Satz 1 ist dieses Integral durch die Funktion $f(x)$ bis auf einen konstanten Summanden eindeutig bestimmt.

---

[1]) Anm. d. Übers.: Vgl. die Fußnote auf S. 205.

Die erhaltenen Resultate lassen sich leicht auf Systeme linearer Differentialgleichungen 1. Ordnung übertragen. Wir beschränken uns hier auf die Angabe der entsprechenden Resultate und lassen die Beweise weg.

Wir betrachten ein homogenes System von n linearen Differentialgleichungen 1. Ordnung

$$y_i' = \sum_{k=1}^{n} a_{ik}(x)\, y_k, \qquad i = 1, 2, \ldots, n, \tag{16}$$

mit beliebig oft differenzierbaren Koeffizienten $a_{ik}(x)$. Ein solches System besitzt eine gewisse Anzahl von „klassischen“ Lösungen (d. h. von Lösungen, die selbst gewöhnliche Funktionen sind und die darüber hinaus noch beliebig oft differenzierbar sind). Man kann zeigen, daß es in der Klasse der Distributionen keine weiteren Lösungen des Systems (16) gibt.

Für das inhomogene System der Form

$$y_i' = \sum_{k=1}^{n} a_{ik} y_k + f_i \tag{17}$$

mit unendlich oft differenzierbaren Koeffizienten a_{ik} und beliebigen Distributionen f_i existiert stets eine Lösung in der Klasse der Distributionen, und diese ist bis auf eine beliebige Lösung des homogenen Systems (16) bestimmt.

Sind im System (17) nicht nur die a_{ik}, sondern auch die f_i „gewöhnliche“ Funktionen, so sind alle in K^* existierenden Lösungen dieses Systems ebenfalls gewöhnliche Funktionen.

4.4.7. Einige Verallgemeinerungen. Bisher haben wir Distributionen einer reellen Veränderlichen betrachtet, d. h. Distributionen auf der reellen Achse. Auf der Grundlage derselben Überlegungen kann man auch Distributionen auf einer beschränkten Menge, etwa auf einem Intervall oder einem Kreis, Distributionen mehrerer Veränderlicher, Distributionen eines komplexen Arguments usw. einführen. Außerdem ist die oben angegebene Definition der Distributionen auf der reellen Achse bei weitem nicht die einzig mögliche. Wir wollen ganz kurz einige der erwähnten Typen von Distributionen betrachten.

a) *Distributionen mehrerer Veränderlicher.* Wir betrachten im n-dimensionalen Raum die Gesamtheit K^n aller Funktionen $\varphi(x_1, x_2, \ldots, x_n)$, die außerhalb eines (jeweils von φ abhängigen) Quaders

$$a_i \leqq x_i \leqq b_i, \qquad i = 1, 2, \ldots, n,$$

verschwinden und für die sämtliche partiellen Ableitungen nach allen Argumenten existieren. Die Menge K^n bildet einen linearen Raum (mit der gewöhnlichen Addition und Skalarmultiplikation von Funktionen). Analog wie im eindimensionalen Fall kann man in diesem Raum eine Konvergenz einführen: φ_k konvergiert gegen φ, $\varphi_k \to \varphi$,

wenn alle Funktionen φ_k außerhalb eines gemeinsamen Quaders $a_i \leqq x_i \leqq b_i$, $i = 1, 2, \ldots, n$, verschwinden und wenn in diesem Quader

$$\frac{\partial^r \varphi_k}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}} \to \frac{\partial^r \varphi}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}} \quad \left(\sum_{i=1}^{n} \alpha_i = r\right)$$

gilt, und zwar gleichmäßig für jedes fixierte n-Tupel $(\alpha_1, \ldots, \alpha_n)$ von nichtnegativen ganzen Zahlen.

Jedes stetige lineare Funktional auf K^n heißt *Distribution von n Veränderlichen.* Jede lokalintegrierbare Funktion $f(x)$ von n Veränderlichen läßt sich gleichzeitig auch als Distribution interpretieren. Der Wert des entsprechenden Funktionals wird durch die Formel

$$(f, \varphi) = \int f(x)\, \varphi(x)\, dx \qquad (x = (x_1, \ldots, x_n),\ dx = dx_1 \cdots dx_n)$$

gegeben. Wie auch im Fall $n = 1$ bestimmen verschiedene stetige Funktionen verschiedene Funktionale (d. h. stellen verschiedene Distributionen dar).[1]

Für Distributionen von n Veränderlichen werden die Begriffe des Grenzüberganges, der Ableitung usw. analog wie im Falle einer Veränderlichen eingeführt. Zum Beispiel werden partielle Ableitungen einer Distribution durch die Formel

$$\left(\frac{\partial^r f(x)}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}, \varphi(x)\right) = (-1)^r \left(f(x), \frac{\partial^r \varphi(x)}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}\right)$$

definiert. Daraus ist ersichtlich, daß jede Distribution von n Veränderlichen partielle Ableitungen beliebiger Ordnung besitzt.

b) *Komplexe Distributionen.* Als Grundfunktionen nehmen wir jetzt die unendlich oft differenzierbaren, finiten, komplexwertigen Funktionen auf der reellen Achse. Die Funktionale auf dem Raum $\tilde{K}$ dieser Grundfunktionen bezeichnen wir dann als *komplexe Distributionen.* Wir erinnern daran, daß in einem komplexen linearen Raum lineare und antilineare Funktionale existieren. Die ersten genügen der Bedingung

$$(f, \alpha\varphi) = \alpha(f, \varphi)$$

(α komplexe Zahl), die zweiten der Bedingung

$$(f, \alpha\varphi) = \bar{\alpha}(f, \varphi).$$

Ist $f(x)$ eine gewöhnliche komplexwertige Funktion auf der reellen Achse, so kann man ihr auf zwei Arten ein lineares Funktional auf $\tilde{K}$ zuordnen:

$$(f, \varphi)_1 = \int_{-\infty}^{\infty} f(x)\, \varphi(x)\, dx \tag{18$_1$}$$

und

$$(f, \varphi)_2 = \int_{-\infty}^{\infty} \overline{f(x)}\, \varphi(x)\, dx. \tag{18$_2$}$$

[1]) Anm. d. Übers.: Die Fußnote auf S. 205 läßt sich sinngemäß übertragen.

Derselben Funktion $f(x)$ kann man auch zwei antilineare Funktionale zuordnen, und zwar

$$_1(f, \varphi) = \int_{-\infty}^{\infty} f(x)\, \overline{\varphi(x)}\, dx \tag{18$_3$}$$

und

$$_2(f, \varphi) = \int_{-\infty}^{\infty} \overline{f(x)}\, \overline{\varphi(x)}\, dx. \tag{18$_4$}$$

Der Wahl einer dieser vier Möglichkeiten entspricht jeweils eine bestimmte Art der Einbettung des Raumes der „gewöhnlichen" Funktionen in den Raum der Distributionen. Die Operationen für komplexe Distributionen werden ganz analog definiert, wie das oben für reelle Distributionen beschrieben wurde.

c) *Distributionen auf einem Kreis.* Manchmal ist es günstig, Distributionen auf einer beschränkten Menge zu betrachten. Als sehr einfaches Beispiel betrachten wir Distributionen auf einem Kreis. Als Raum der Grundfunktionen nehmen wir die Gesamtheit aller unendlich oft differenzierbaren Funktionen auf dem Kreis, dabei sollen die Addition und Skalarmultiplikation in der üblichen Weise erklärt sein. Eine Folge $\{\varphi_n(x)\}$ von Funktionen aus diesem Raum nennen wir konvergent, wenn für jedes $k = 0, 1, 2, \ldots$ die Folge der Ableitungen $\{\varphi_n^{(k)}(x)\}$ auf dem ganzen Kreis gleichmäßig konvergiert. Da hier die Menge der Argumente (der Kreis) beschränkt ist, entfällt die Forderung der Finitheit der Grundfunktionen automatisch. Lineare Funktionale auf diesem Raum nennen wir *Distributionen auf dem Kreis.*

Jede gewöhnliche Funktion auf dem Kreis kann man auch als periodische Funktion ansehen, die auf der reellen Achse gegeben ist. Wenn man diese Auffassung auf die Distributionen bezieht, kann man Distributionen auf einem Kreis auch als periodische Distributionen ansehen. Dabei ist unter einer *periodischen Distribution* (mit der Periode a) natürlich ein Funktional f zu verstehen, das der Bedingung

$$\big(f(x), \varphi(x - a)\big) = \big(f(x), \varphi(x)\big)$$

für jede Grundfunktion φ genügt. Als Beispiel einer periodischen Distribution kann die Distribution

$$\sum_{n=1}^{\infty} \cos nx = -\frac{1}{2} + \pi \sum_{k=-\infty}^{\infty} \delta(x - 2k\pi)$$

dienen, die bereits oben erwähnt wurde.

d) *Andere Grundräume.* Oben haben wir Distributionen auf der reellen Achse als lineare Funktionale auf dem Raum K der unendlich oft differenzierbaren finiten Funktionen definiert. Jedoch die Wahl dieses Grundraumes ist nicht die einzig mögliche. Zum Beispiel könnte man anstelle des Raumes K der finiten Funktionen den größeren Raum S_∞ nehmen. Dieser Raum besteht aus allen beliebig oft differenzierbaren Funktionen $\varphi(x)$ auf der reellen Achse, die zusammen mit ihren Ableitun-

gen schneller als jede Potenz von $1/|x|$ fallen. Genauer gesagt, soll eine Funktion $\varphi(x)$ zum Grundraum S_∞ gehören, wenn für je zwei Zahlen $p, q = 0, 1, 2, \ldots$ eine Konstante $C_{p,q}$ (die von p, q und φ abhängt) mit

$$|x^p \varphi^{(q)}(x)| < C_{p,q}, \qquad -\infty < x < \infty, \tag{19}$$

existiert. Die Konvergenz in S_∞ wird folgendermaßen erklärt: Eine Folge $\{\varphi_n(x)\}$ heißt konvergent gegen $\varphi(x)$, wenn die Folge $\{\varphi_n{}^{(q)}(x)\}$ für jedes $q = 0, 1, \ldots$ auf jedem endlichen Intervall gleichmäßig gegen $\varphi(x)$ konvergiert und wenn man die Konstanten $C_{p,q}$ in den Ungleichungen

$$|x^p \varphi_n{}^{(q)}(x)| < C_{p,q}$$

unabhängig von n wählen kann. Dabei erhält man einen etwas kleineren Raum von Distributionen als im Fall des Raumes K. Zum Beispiel definiert die Funktion

$$f(x) = e^{x^2}$$

ein stetiges lineares Funktional auf K, aber nicht auf S_∞. Die Wahl von S_∞ als Grundraum ist zum Beispiel bei der Betrachtung der Fouriertransformation von Distributionen zweckmäßig.

Wie die Entwicklung der Theorie der Distributionen zeigte, ist es allgemein nicht notwendig, sich ein für allemal auf einen bestimmten Grundraum festzulegen. Vielmehr ist es günstig, ihn zweckentsprechend in Abhängigkeit vom betrachteten Problemkreis zu variieren. Dabei besteht die wesentliche Forderung allein darin, daß es einerseits „hinreichend viele" Grundfunktionen gibt (um mit ihrer Hilfe die „gewöhnlichen" Funktionen oder genauer die regulären Funktionale zu trennen) und daß andererseits diese Grundfunktionen hinreichend glatt sind.

Aufgabe. Man beweise, daß man den Raum S_∞ zu einem abzählbar-normierten Raum machen kann, wenn man etwa

$$\|\varphi\|_n = \sum_{p+q=n} \sup_{\substack{-\infty<x<\infty \\ 0 \le i \le p \\ 0 \le j \le q}} |(1 + |x|^i)\, \varphi^{(j)}(x)|$$

setzt. Ferner zeige man, daß eine im obigen Sinne konvergente Folge auch bezüglich der Topologie konvergiert, die durch diese Normen definiert wird.

4.5. Lineare Operatoren

4.5.1. Definition und Beispiele linearer Operatoren. Es seien E und E_1 zwei lineare topologische Räume. Eine Abbildung

$$y = Ax \qquad (x \in E,\ y \in E_1)$$

aus E in E_1, die der Bedingung

$$A(\alpha x_1 + \beta x_2) = \alpha A x_1 + \beta A x_2$$

genügt, nennen wir *linearen Operator*. Die Gesamtheit D_A aller Elemente $x \in E$, für die die Abbildung A definiert ist, heißt *Definitionsbereich* von A. Im allgemeinen setzen wir nicht voraus, daß $D_A = E$ ist. Jedoch werden wir stets annehmen, daß D_A eine lineare Menge ist, d. h., daß mit $x, y \in D_A$ auch $\alpha x + \beta y \in D_A$ ist für alle α, β.

Ein Operator A heißt *stetig im Punkt* $x_0 \in D_A$, wenn zu jeder Umgebung V des Punktes $y_0 = Ax_0$ eine Umgebung U des Punktes x_0 existiert, so daß $Ax \in V$ für alle $x \in U \cap D_A$ ist. Ein Operator A heißt *stetig*, wenn er in jedem Punkt $x \in D_A$ stetig ist. Sind E und E_1 normierte Räume, so ist diese Definition mit der folgenden gleichwertig: Ein Operator A heißt stetig, wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so daß

$$\|Ax' - Ax''\| < \varepsilon$$

ist für alle $x', x'' \in D_A$ mit

$$\|x' - x''\| < \delta.$$

Der Begriff des linearen Funktionals, der am Anfang dieses Kapitels eingeführt wurde, ist ein Spezialfall eines linearen Operators. Und zwar ist ein lineares Funktional ein linearer Operator, der den gegebenen Raum E in die reellen Zahlen R^1 abbildet. Die Definitionen der Linearität und der Stetigkeit eines Operators gehen für $E_1 = R^1$ in die entsprechenden Definitionen über, die früher für Funktionale gegeben wurden.

Beispiele linearer Operatoren

1. E sei ein topologischer linearer Raum, und es sei

$$Ax = x \quad \text{für alle } x \in R.$$

Ein solcher Operator, der jedes Element des Raumes in sich überführt, heißt *identischer Operator*.

2. E_1 und E_2 seien beliebige topologische lineare Räume, und es sei

$$0x = 0 \quad \text{für alle } x \in E.$$

(Hier ist 0 das Nullelement des Raumes E_1.) Dann heißt 0 *Nulloperator*.

3. *Die allgemeine Gestalt eines linearen Operators, der einen endlichdimensionalen Raum in einen endlichdimensionalen Raum überführt.* A sei ein linearer Operator, der den n-dimensionalen Raum $\boldsymbol{R}^n$ mit der Basis $e_1, \ldots, e_n$ in den m-dimensionalen Raum $\boldsymbol{R}^m$ mit der Basis $f_1, \ldots, f_m$ überführt. Für einen beliebigen Vektor x des $\boldsymbol{R}^n$ gilt dann

$$x = \sum_{i=1}^{n} x_i e_i$$

und wegen der Linearität des Operators A

$$Ax = \sum_{i=1}^{n} x_i A e_i.$$

Somit ist der Operator A bestimmt, wenn man weiß, in welchen Raum er die Basisvektoren $e_1, \ldots, e_n$ abbildet. Wir stellen die Vektoren Ae_i als Linearkombination der Basisvektoren $f_1, \ldots, f_m$ dar,

$$Ae_i = \sum_{k=1}^{m} a_{ki} f_k .$$

Hieraus ist ersichtlich, daß der Operator A durch die Koeffizientenmatrix $\|a_{ki}\|$ bestimmt ist. Das Bild des $\boldsymbol{R}^n$ im Raum $\boldsymbol{R}^m$ bildet dort einen linearen Teilraum, dessen Dimension offenbar gleich dem Rang der Matrix $\|a_{ki}\|$, also keinesfalls größer als n ist. Wir bemerken, daß jeder lineare Operator auf einem endlichdimensionalen Raum automatisch stetig ist.

4. H sei ein Hilbertraum und H_1 ein Teilraum von H. Wir zerlegen H in die direkte Summe des Teilraumes H_1 und seines orthogonalen Komplements, d. h., wir stellen jedes Element $h \in H$ in der Form

$$h = h_1 + h_2 \qquad (h_1 \in H_1, h_2 \perp H_1)$$

dar. Da diese Darstellung eindeutig ist, können wir

$$Ph = h_1$$

setzen. Diesen Operator P bezeichnet man üblicherweise als *orthogonalen Projektor* oder kurz *Orthoprojektor* von H auf H_1. Die Linearität und die Stetigkeit prüft man mühelos nach.

5. Im Raum der stetigen Funktionen auf dem Intervall $[a, b]$ betrachten wir den durch die Formel

$$\psi(s) = \int_a^b K(s, t)\, \varphi(t)\, dt \tag{1}$$

definierten Operator, $K(s, t)$ ist dabei eine fixierte stetige Funktion zweier Veränderlicher. Die Funktion $\psi(t)$ ist für jede stetige Funktion $\varphi(s)$ wieder stetig, so daß der Operator (1) den Raum der stetigen Funktionen tatsächlich in sich überführt. Die Linearität des Operators (1) ist trivial. Um von seiner Stetigkeit reden zu können, muß man zunächst erst einmal angeben, welche Topologie in diesem Funktionenraum betrachtet wird. Dem Leser wird empfohlen, die Stetigkeit des Operators zu beweisen, wenn a) der Raum $C[a, b]$, d. h. der Raum der stetigen Funktionen mit der Norm $\|\varphi\| = \max |\varphi(t)|$, und wenn b) der Raum $C^2[a, b]$ betrachtet wird, d. h., wenn $\|\varphi\| = \left(\int_a^b \varphi^2(t)\, dt \right)^{1/2}$ ist.

6. In demselben Raum der stetigen Funktionen betrachten wir den Operator

$$\psi(t) = \varphi_0(t)\, \varphi(t),$$

wobei $\varphi_0(t)$ eine fixierte stetige Funktion ist. Die Linearität dieses Operators ist trivial. (Man zeige seine Stetigkeit bei den im vorhergehenden Beispiel angegebenen Normierungen.)

7. Eines der wichtigsten Beispiele von linearen Operatoren bildet der Differentiationsoperator. Man kann ihn in verschiedenen Räumen betrachten.

a) Zunächst betrachten wir den Operator

$$Df(t) = f'(t)$$

im Raum $C[a, b]$ der stetigen Funktionen. Dieser Operator (den wir als Abbildung aus $C[a, b]$ in $C[a, b]$ ansehen) ist offenbar nicht auf dem ganzen Raum der stetigen Funktionen definiert, sondern nur auf der linearen Menge aller stetig differenzierbaren Funktionen. Der Operator D ist linear, aber nicht stetig. Das ist beispielsweise daraus ersichtlich, daß die Folge

$$\varphi_n(t) = \frac{\sin nt}{t}$$

(in der Metrik des Raumes $C[a, b]$) gegen 0 konvergiert, während die Folge

$$D\varphi_n(t) = \cos nt$$

nicht konvergiert.

b) Den Differentiationsoperator kann man auch als Operator aus dem Raum D_1 der auf $[a, b]$ stetig differenzierbaren Funktionen mit der Norm

$$\|\varphi\|_1 = \max|\varphi(t)| + \max|\varphi'(t)|$$

in den Raum $C[a, b]$ auffassen. In diesem Fall ist der Operator D linear und stetig und bildet ganz D_1 auf ganz $C[a, b]$ ab.

c) Den Differentiationsoperator als Operator aus D_1 in $C[a, b]$ aufzufassen ist nicht immer zweckmäßig, obwohl wir damit einen stetigen linearen Operator erhalten, der auf dem ganzen Raum definiert ist. Denn man kann diesen Operator im allgemeinen nicht zweimal auf eine Funktion aus D_1 anwenden. Dazu ist es günstiger, den Differentiationsoperator in einem noch kleineren Raum als D_1 zu betrachten, und zwar im Raum D_∞ der unendlich oft differenzierbaren Funktionen auf dem Intervall $[a, b]$. Die Topologie in D_∞ erklären wir dabei durch das abzählbare Normensystem

$$\|\varphi\|_n = \sup_{n,t}\left(|\varphi(t)|, |\varphi'(t)|, \ldots, |\varphi^{(n)}(t)|\right).$$

Der Differentiationsoperator führt den ganzen Raum D_∞ in sich über und ist, wie man leicht nachprüfen kann, stetig auf diesem Raum.

d) Die unendlich oft differenzierbaren Funktionen wiederum sind eine sehr enge Klasse. Eine Möglichkeit, den Differentiationsoperator in einem wesentlich größeren Raum zu betrachten, so daß er dabei aber noch stetig bleibt, bieten die Distribu-

tionen. In 4.4. haben wir bereits gesagt, wie die Differentiation von Distributionen erklärt ist. Aus dem dort Gesagten folgt offenbar, daß die Differentiation ein linearer Operator im Raum der Distributionen ist. Dieser Operator ist in dem Sinne stetig, daß aus der Konvergenz einer Folge $\{f_n(t)\}$ von Distributionen gegen eine Distribution $f(t)$ die Konvergenz der Folge aller Ableitungen gegen die entsprechenden Ableitungen von $f(t)$ folgt.

4.5.2. Stetigkeit und Beschränktheit. Ein linearer Operator aus E in E_1 heißt *beschränkt,* wenn er auf ganz E definiert ist und jede beschränkte Menge wieder in eine beschränkte Menge überführt. Zwischen der Beschränktheit und der Stetigkeit eines linearen Operators besteht ein enger Zusammenhang, und zwar gilt folgende Behauptung.

I. *Jeder stetige lineare Operator ist beschränkt.*

Zum Beweis nehmen wir an, daß eine beschränkte Menge $M \subset E$ existiert, deren Bild $AM \subset E_1$ nicht beschränkt ist. Dann findet man in E_1 eine Nullumgebung V, so daß keine der Mengen AM/n in V enthalten ist. Dann existiert aber eine Folge $x_n \in M$, für die keines der Elemente Ax_n/n zu V gehört. Die Folge $\{Ax_n/n\}$ konvergiert also nicht gegen Null in E_1. Andererseits gilt $x_n/n \to 0$ in E,[1]) und das widerspricht der Stetigkeit des Operators A.

II. *Ist A ein beschränkter linearer Operator von E in E_1 und genügt E dem ersten Abzählbarkeitsaxiom, so ist A stetig.*

Zum Beweis nehmen wir an, daß A nicht stetig ist. Dann gibt es eine Nullumgebung V in E_1 und eine Nullumgebungsbasis $\{U_n\}$ in E, so daß für jedes n ein Element $x_n \in U_n/n$ mit $Ax_n \notin nV$ existiert. Die Folge x_n ist dann in E beschränkt (sie strebt sogar gegen 0), während die Folge Ax_n in E_1 nicht beschränkt ist (da sie in keiner der Mengen nV enthalten ist). Also kann A nicht beschränkt sein. Damit haben wir die Kontraposition der Behauptung bewiesen.

Also ist in Räumen, die dem ersten Abzählbarkeitsaxiom genügen (zu denen insbesondere alle normierten und abzählbar-normierten Räume gehören), *die Stetigkeit eines Operator gleichbedeutend mit seiner Beschränktheit.*

Alle Operatoren, die in den Beispielen 1 bis 6 des vorhergehenden Abschnitts angeführt wurden, sind stetig. Da alle dort betrachteten Räume dem ersten Abzählbarkeitsaxiom genügen, sind alle dort aufgezählten Operatoren auch beschränkt.

Sind E und E_1 normierte Räume, so kann man die Beschränktheit eines Operators A von E in E_1 auch folgendermaßen formulieren: Ein Operator A heißt *beschränkt,* wenn er jede Kugel in eine beschränkte Menge überführt. Auf Grund der Linearität kann diese Bedingung wiederum so formuliert werden: Ein Operator A ist beschränkt, wenn eine Konstante C existiert, so daß für alle $f \in E$

$$\|Af\| \leqq C\,\|f\|$$

[1]) Vgl. 3.5.1., Aufgabe 1.

ist. Die kleinste Zahl C, die dieser Ungleichung genügt, heißt *Norm* des Operators A und wird mit $\|A\|$ bezeichnet.

Satz 1. *Für jeden beschränkten Operator A, der einen normierten Raum in einen normierten Raum abbildet, gilt*

$$\|A\| = \sup_{\|x\| \leqq 1} \|Ax\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}. \tag{2}$$

Beweis. Setzen wir $\alpha = \sup\limits_{\|x\| \leqq 1} \|Ax\|$, so gilt wegen der Linearität von A die Gleichheit

$$\alpha = \sup_{\|x\| \leqq 1} \|Ax\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}.$$

Daher ist für jedes Element $x \neq 0$

$$\frac{\|Ax\|}{\|x\|} \leqq \alpha$$

bzw.

$$\|Ax\| \leqq \alpha \, \|x\|.$$

Daraus ergibt sich

$$\|A\| = \inf C \leqq \alpha.$$

Andererseits existiert zu jedem $\varepsilon > 0$ ein Element $x_\varepsilon \neq 0$ mit

$$\alpha - \varepsilon \leqq \frac{\|Ax_\varepsilon\|}{\|x_\varepsilon\|}$$

bzw.

$$(\alpha - \varepsilon) \, \|x_\varepsilon\| \leqq \|Ax_\varepsilon\| \leqq C \, \|x_\varepsilon\|.$$

Daher gilt

$$\alpha - \varepsilon \leqq \inf C = \|A\|$$

und damit auch

$$\alpha \leqq \|A\|,$$

denn ε war beliebig. Folglich ist $\|A\| = \alpha$.

4.5.3. Summe und Produkt von Operatoren

Definition 1. A und B seien zwei lineare Operatoren, die aus dem linearen Raum E in den linearen Raum E_1 abbilden. Als ihre *Summe* $A + B$ bezeichnen wir den Operator C, der dem Element $x \in E$ das Element

$$y = Ax + Bx \in E_1$$

zuordnet. C ist für alle Elemente aus dem Durchschnitt $D_A \cap D_B$ der Definitionsgebiete von A bzw. B definiert.

Man prüft leicht nach, daß $C = A + B$ ein linearer Operator ist und daß C stetig ist, falls A und B stetig waren.

Sind E und E_1 normierte Räume und die Operatoren A und B beschränkt, so ist auch $A + B$ beschränkt; dabei gilt

$$\|A + B\| \leqq \|A\| + \|B\|. \tag{3}$$

Das ergibt sich aus der für jedes x gültigen Beziehung

$$\|(A + B)\, x\| = \|Ax + Bx\| \leqq \|Ax\| + \|Bx\| \\ \leqq (\|A\| + \|B\|)\, \|x\|.$$

Definition 2. A sei ein linearer Operator, der aus dem Raum E in den Raum E_1 abbildet, und B sei ein linearer Operator, der aus E_1 in E_2 abbildet. Als *Produkt* BA der Operatoren A und B wird der Operator C bezeichnet, der dem Element $x \in E$ das Element

$$z = B(Ax)$$

aus E_2 zuordnet. Der Definitionsbereich D_C des Operators $C = BA$ besteht aus allen $x \in D_A$, für die $Ax \in D_B$ gilt.

Offenbar ist der Operator BA linear. Er ist auch stetig, wenn A und B stetig sind.

Aufgabe. Es ist zu beweisen, daß D_C eine lineare Menge ist, wenn D_A und D_B lineare Mengen sind.

Sind A und B beschränkte Operatoren zwischen normierten Räumen, so ist auch der Operator BA beschränkt, und dabei gilt

$$\|BA\| \leqq \|B\| \cdot \|A\|. \tag{4}$$

Das ergibt sich aus der Beziehung

$$\|B(Ax)\| \leqq \|B\| \cdot \|Ax\| \leqq \|B\| \cdot \|A\| \cdot \|x\|. \tag{5}$$

Die Summe und das Produkt von drei und mehr Operatoren werden induktiv definiert. Beide Operationen sind assoziativ.

Als Produkt kA eines Operators A mit der Zahl k wird der Operator definiert, der jedem Element x das Element kAx zuordnet.

Die Gesamtheit $\mathscr{L}(E, E_1)$ aller stetigen linearen Operatoren von E in E_1 (E und E_1 fixierte topologische lineare Räume) bilden bezüglich der oben eingeführten Addition und Skalarmultiplikation einen linearen Raum. Sind E und E_1 normierte Räume, so ist auch $\mathscr{L}(E, E_1)$ ein normierter Raum (mit der oben angegebenen Definition der Norm eines Operators).

Aufgabe. E sei ein normierter und E_1 ein vollständiger normierter Raum. Dann gelten die folgenden beiden Aussagen:

a) Der normierte Raum $\mathcal{L}(E, E_1)$ ist vollständig.

b) Ist $A_k \in \mathcal{L}(E, E_1)$ und $\sum_{k=1}^{\infty} \|A_k\| < \infty$, so konvergiert die Reihe $\sum_{k=1}^{\infty} A_k$ gegen einen Operator $A \in \mathcal{L}(E, E_1)$, und es ist

$$\|A\| = \left\| \sum_{k=1}^{\infty} A_k \right\| \leqq \sum_{k=1}^{\infty} \|A_k\| . \tag{6}$$

4.5.4. Der Umkehroperator, Umkehrbarkeit. Es sei A ein Operator aus E in E_1, D_A der Definitionsbereich von A und R_A der Wertebereich von A.

Definition 3. Der Operator A heißt *umkehrbar* (*invertierbar*), wenn die Gleichung

$$Ax = y$$

für jedes $y \in R_A$ genau eine Lösung besitzt.

Ist A umkehrbar, so kann man jedem $y \in R_A$ in eindeutiger Weise ein Element $x \in D_A$ zuordnen, und zwar die Lösung der Gleichung $Ax = y$. Der Operator, der diese Zuordnung vermittelt, heißt *Umkehroperator* von A (*inverser Operator* von A) und wird mit A^{-1} bezeichnet.

Satz 2. *Der Umkehroperator A^{-1} eines linearen Operators A ist ebenfalls linear.*

Beweis. Wir stellen zunächst fest, daß DA^{-1}, d. h. der Wertebereich des Operators A, eine lineare Menge ist. Sind $y_1, y_2 \in R_A$, so müssen wir die Beziehung

$$A^{-1}(\alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 A^{-1} y_1 + \alpha_2 A^{-1} y_2 \tag{7}$$

nachweisen. Dazu setzen wir $y_1 = Ax_1$ und $y_2 = Ax_2$, dann gilt wegen der Linearität von A

$$A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 y_1 + \alpha_2 y_2 . \tag{8}$$

Nach Definition des Umkehroperators ist

$$A^{-1} y_1 = x_1 , \qquad A^{-1} y_2 = x_2 .$$

Daraus ergibt sich nach Multiplikation dieser Gleichungen mit α_1 bzw. α_2 und anschließender Addition

$$\alpha_1 A^{-1} y_1 + \alpha_2 A^{-1} y_2 = \alpha_1 x_1 + \alpha_2 x_2 .$$

Andererseits folgt aus (8) und aus der Definition des Umkehroperators

$$\alpha_1 x_1 + \alpha_2 x_2 = A^{-1}(\alpha_1 y_1 + \alpha_2 y_2) .$$

Das ergibt zusammen mit der vorhergehenden Gleichung

$$A^{-1}(\alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 A^{-1} y_1 + \alpha_2 A^{-1} y_2 .$$

Satz 3 (Satz von BANACH über den Umkehroperator). *A sei ein eineindeutiger beschränkter linearer Operator des Banachraumes E auf den Banachraum E_1. Dann ist der Umkehroperator A^{-1} beschränkt.*

Zum Beweis benötigen wir folgendes Lemma.

Lemma. *M sei eine überall dichte Teilmenge eines Banachraumes E. Dann kann man jedes von Null verschiedene Element $y \in E$ in eine Reihe*

$$y = y_1 + y_2 + \cdots + y_n + \cdots$$

mit $y_k \in M$ und $\|y_k\| \leqq \dfrac{3\,\|y\|}{2^k}$ entwickeln.

Beweis. Die Elemente y_k konstruieren wir sukzessiv. Wir wählen zunächst ein y_1 mit

$$\|y - y_1\| \leqq \frac{\|y\|}{2}. \tag{9}$$

Das ist möglich, da Ungleichung (9) eine Kugel vom Radius $\dfrac{\|y\|}{2}$ um den Punkt y beschreibt, in der es ein Element aus M geben muß (M war überall dicht in E). Nun wählen wir ein $y_2 \in M$ mit $\|y - y_1 - y_2\| \leqq \|y\|/4$, anschließend ein y_3 mit $\|y - y_1 - y_2 - y_3\| \leqq \|y\|/8$ usw. Allgemein wählen wir y_n so, daß

$$\|y - y_1 - \cdots - y_n\| \leqq \|y\|/2^n$$

ist. Eine solche Wahl ist stets möglich, da M überall dicht ist in E. Auf Grund der Wahl der Elemente y_k gilt

$$\left\| y - \sum_{k=1}^{n} y_k \right\| \to 0 \quad \text{für} \quad n \to \infty,$$

d. h., die Reihe $\sum\limits_{k=1}^{\infty} y_k$ konvergiert gegen y. Wir schätzen noch die Normen der Elemente y_k ab, es ist

$$\|y_1\| = \|y_1 - y + y\| \leqq \|y_1 - y\| + \|y\| \leqq 3\,\|y\|/2,$$
$$\|y_2\| = \|y_2 + y_1 - y + y - y_1\| \leqq \|y - y_1 - y_2\| + \|y - y_1\| \leqq 3\,\|y\|/4$$

und schließlich

$$\|y_n\| = \|y_n + y_{n-1} + \cdots + y_1 - y + y - y_1 - \cdots - y_{n-1}\|$$
$$\leqq \|y - y_1 - \cdots - y_n\| + \|y - y_1 - \cdots - y_{n-1}\| \leqq \frac{3\,\|y\|}{2^n}.$$

Damit ist das Lemma bewiesen.

Beweis von Satz 3. Wir betrachten im Raum E_1 die Menge M_k aller Elemente y mit $\|A^{-1}y\| \leqq k\,\|y\|$. Jedes Element des Raumes E_1 liegt dann in einer Menge M_k,

d. h., es ist $E_1 = \bigcup_{k=1}^{\infty} M_k$. Nach dem Satz von BAIRE (2.3.3.) ist wenigstens eine der Mengen M_k, etwa M_n, dicht in einer gewissen Kugel B. Im Innern der Kugel B wählen wir einen Kreisring, in dessen Zentrum ein Punkt aus M_n liegt. Dieser Kreisring P sei die Gesamtheit aller Punkte z mit $\beta < \|z - y_0\| < \alpha$, $0 < \beta < \alpha$, $y_0 \in M_n$. Wenn wir den Kreisring P so verschieben, daß sein Zentrum mit dem Koordinatenanfangspunkt zusammenfällt, erhalten wir den Kreisring $P_0 = \{z : 0 < \beta < \|z\| < \alpha\}$.

Wir wollen zeigen, daß in P_0 eine gewisse Menge M_N dicht liegt. Dazu sei $z \in P \cap M_n$, dann ist $z - y_0 \in P_0$ und

$$\begin{aligned}\|A^{-1}(z - y_0)\| &\leqq \|A^{-1}z\| + \|A^{-1}y_0\| \leqq n(\|z\| + \|y_0\|)\\ &\leqq n(\|z - y_0\| + 2\,\|y_0\|) = n\,\|z - y_0\| \left(1 + \frac{2\,\|y_0\|}{\|z - y_0\|}\right)\\ &\leqq n\,\|z - y_0\| \left(1 + \frac{2\,\|y_0\|}{\beta}\right).\end{aligned} \tag{10}$$

Die Größe $n\left(1 + \frac{2\,\|y_0\|}{\beta}\right)$ hängt dabei nicht von z ab. Setzen wir[1])

$$N = 1 + n\left[1 + \frac{2\,\|y_0\|}{\beta}\right],$$

so gilt $z - y_0 \in M_N$ auf Grund von (10). Da M_n dicht in P ist, liegt M_N folglich dicht in P_0.

Wir zeigen jetzt, daß M_N sogar in E_1 dicht liegt. Es sei $y \in E_1$ von Null verschieden. Dann kann man λ stets so wählen, daß $\beta < \|\lambda y\| < \alpha$ ist, d. h. λy zu P_0 gehört. Da M_N in P_0 dicht liegt, gibt es eine Folge $y_k \in M_N$, die gegen λy konvergiert. Die Folge y_k/λ konvergiert dann gegen y. Da y_k zu M_N gehört, liegt offenbar auch y_k/λ in M_N für jede reelle Zahl $\lambda \neq 0$. Somit ist M_N dicht in $E_1 \setminus \{0\}$ und folglich auch in E_1.

Ist $y \in E_1$ ein von Null verschiedenes Element, so kann man es nach dem gezeigten Lemma in eine Reihe

$$y = y_1 + y_2 + \cdots + y_k + \cdots$$

nach Elementen aus M_N entwickeln, wobei $\|y_k\| \leqq \frac{3\,\|y\|}{2^k}$ ist. Die Reihe der Urbilder $x_k = A^{-1}y_k$ der Elemente y_k konvergiert im Raum E auf Grund der Ungleichung

$$\|x_k\| = \|A^{-1}y_k\| \leqq N\,\|y_k\| < N\,\frac{3\,\|y\|}{2^k}$$

gegen ein gewisses Element $x \in E$. Dabei gilt

$$\|x\| \leqq \sum_{k=1}^{\infty} \|x_k\| \leqq 3N\,\|y\| \sum_{k=1}^{\infty} \frac{1}{2^k} = 3N\,\|y\|.$$

---

[1]) Die Klammern [ ] bedeuten den ganzen Teil der Zahl.

Wegen der Konvergenz der Reihe $\sum_{n=1}^{\infty} x_n$ und der Stetigkeit des Operators A ist

$$Ax = A\left(\sum_{n=1}^{\infty} x_n\right) = \sum_{n=1}^{\infty} Ax_n = \sum_{n=1}^{\infty} y_n = y,$$

woraus $x = A^{-1}y$ folgt. Dabei gilt

$$\|A^{-1}y\| = \|x\| \leqq 3N\,\|y\|,$$

und da diese Abschätzung für jedes $y \neq 0$ richtig ist, muß A^{-1} beschränkt sein. Damit ist der Satz bewiesen.

Aufgaben

1. E und E_1 seien normierte Räume. Ein linearer Operator aus E in E_1 (dessen Definitionsgebiet $D_A \subset E$ eine lineare Menge ist) heißt *abgeschlossen*, wenn aus den Bedingungen $x_n \in D_A$, $x_n \to x$, $Ax_n \to y$ folgt, daß $x \in D_A$ und $Ax = y$ ist. Man zeige, daß jeder beschränkte Operator abgeschlossen ist.

2. Wir betrachten das direkte Produkt $E \times E_1$ der Räume E und E_1. Das ist der lineare normierte Raum aller Paare $[x, y]$, $x \in E$, $y \in E_1$, mit der Norm $\|[x, y]\| = \|x\| + \|y\|_1$ ($\|\cdot\|$ und $\|\cdot\|_1$ sind die Normen in E bzw. E_1). Einem Operator A kann man dann die Menge $G_A = \{[x, y]: x \in D_A, y = Ax\} \subset E \times E_1$, den sogenannten *Graphen* von A, zuordnen. Man zeige, daß G_A eine lineare Menge in $E \times E_1$ bildet, die genau dann abgeschlossen ist, wenn der Operator A abgeschlossen ist. Weiter zeige man, daß *im Fall von Banachräumen E und E_1 ein auf ganz E definierter abgeschlossener Operator auch beschränkt ist* (Satz vom abgeschlossenen Graphen).

Hinweis. Man wende Satz 3 auf den Operator $P: [x, Ax] \to x$ an, der G_A in E abbildet.

3. E und E_1 seien vollständige abzählbar-normierte Räume, und A sei ein stetiger linearer Operator, der E auf E_1 abbildet. Man beweise, daß dann der Umkehroperator A^{-1} stetig ist. Ferner formuliere und beweise man den Satz vom abgeschlossenen Graphen für abzählbar-normierte Räume.

Wir betrachten die Menge $\mathscr{L}(E, E_1)$ aller beschränkten linearen Operatoren A, die den Banachraum E in den Banachraum E_1 abbilden, $\mathscr{L}(E, E_1)$ ist ebenfalls ein Banachraum. Mit $\mathscr{G}\mathscr{L}(E, E_1)$ bezeichnen wir die Menge aller Operatoren aus $\mathscr{L}(E, E_1)$, die E auf ganz E_1 abbilden und einen beschränkten Umkehroperator besitzen. Diese Menge ist offen in $\mathscr{L}(E, E_1)$. Es gilt folgender Satz.

Satz 4. *Es sei $A_0 \in \mathscr{G}\mathscr{L}(E, E_1)$ und ΔA ein beliebiger Operator aus $\mathscr{L}(E, E_1)$ mit $\|\Delta A\| < \dfrac{1}{\|A_0^{-1}\|}$. Dann existiert der Operator $(A_0 + \Delta A)^{-1}$ und ist beschränkt, d. h., $A = A_0 + \Delta A$ gehört zu $\mathscr{G}\mathscr{L}(E, E_1)$.*

Beweis. Wir halten $y \in E_1$ fest und betrachten die durch die Formel

$$Bx = A_0^{-1}y - A_0^{-1}\,\Delta A x$$

definierte Abbildung B des Raumes E in sich. Aus der Bedingung $\|\Delta A\| < \|A_0^{-1}\|^{-1}$ folgt, daß B eine Kontraktion ist. Wegen der Vollständigkeit von E existiert dann genau ein Fixpunkt x der Abbildung B,

$$x = Bx = A_0^{-1}y - A_0^{-1}\,\Delta A x,$$

woraus sich

$$Ax = A_0x + \Delta Ax = y$$

ergibt. Ist $Ax' = y$, so ist x' ebenfalls ein Fixpunkt der Abbildung B, und daher folgt $x = x'$. Somit hat die Gleichung $Ax = y$ für jedes $y \in E_1$ genau eine Lösung. Also besitzt A einen Umkehroperator A^{-1}, der auf ganz E_1 definiert ist. Nach Satz 3 ist der Operator A^{-1} beschränkt, was zu beweisen war.

Satz 5. *Es sei E ein Banachraum, I der identische Operator auf E und A ein beschränkter linearer Operator von E in sich mit $\|A\| < 1$. Dann existiert der Operator $(I - A)^{-1}$, er ist beschränkt und in der Form*

$$(I - A)^{-1} = \sum_{k=0}^{\infty} A^k. \tag{11}$$

darstellbar.

Beweis. Die Existenz und Beschränktheit des Operators $(I - A)^{-1}$ ergibt sich aus Satz 4 (außerdem folgt das auch aus den unten angeführten Überlegungen).

Wegen $\|A\| < 1$ ist $\sum_{k=0}^{\infty} \|A^k\| \leqq \sum_{k=0}^{\infty} \|A\|^k < \infty$. Der Raum E ist vollständig, daher folgt aus der Konvergenz der Reihe $\sum_{k=0}^{\infty} \|A^k\|$, daß die Summe der Reihe $\sum_{k=0}^{\infty} A^k$ ein beschränkter linearer Operator ist. Für jedes n gilt nun

$$(I - A) \sum_{k=0}^{n} A^k = \sum_{k=0}^{n} A^k(I - A) = I - A^{n+1}.$$

Daraus erhalten wir durch den Grenzübergang $n \to \infty$

$$(I - A) \sum_{k=0}^{\infty} A^k = \sum_{k=0}^{\infty} A^k(I - A) = I,$$

dabei haben wir $\|A^{n+1}\| \leqq \|A\|^{n+1} \to 0$ ausgenutzt. Hieraus folgt aber

$$(I - A)^{-1} = \sum_{k=0}^{\infty} A^k,$$

was zu zeigen war.

Aufgabe. A sei ein beschränkter linearer Operator des Banachraumes E auf den Banachraum E_1. Man beweise, daß eine Konstante $\alpha > 0$ existiert, so daß alle Operatoren $B \in \mathscr{L}(E, E_1)$ mit $\|A - B\| < \alpha$ den Raum E auf ganz E_1 abbilden (Banach).

4.5.5. Adjungierte Operatoren. Es seien E und E_1 zwei topologische lineare Räume, und A sei ein stetiger linearer Operator von E in E_1. Weiter sei g ein stetiges lineares Funktional auf E_1, d. h. $g \in E_1^*$. Wie man leicht nachprüfen kann, ergibt die Anwendung des Funktionals g auf die Elemente $y = Ax$ ein stetiges lineares Funktional $f(x) = g(Ax)$ auf E. Das Funktional f ist somit ein Element des Raumes E^*. Jedem

Funktional $g \in E_1^*$ haben wir also ein Funktional $f \in E^*$ zugeordnet, d. h., wir haben einen Operator erhalten, der E_1^* in E^* abbildet. Dieser heißt *adjungierter Operator* des Operators A und wird mit A^* bezeichnet.

Wenn wir den Wert eines Funktionals f auf einem Element x mit dem Symbol (f, x) bezeichnen, erhalten wir $(g, Ax) = (f, x)$ oder

$$(g, Ax) = (A^*g, x).$$

Diese Gleichung kann man als Definitionsgleichung des adjungierten Operators ansehen.

Beispiel. *Der adjungierte Operator im endlichdimensionalen Fall.* Der Operator A möge den n-dimensionalen Raum $\boldsymbol{R}^n$ in dem m-dimensionalen Raum $\boldsymbol{R}^m$ abbilden, und (a_{ij}) sei die Matrix von A.

Die Abbildung $y = Ax$ kann man in Form des Gleichungssystems

$$y_i = \sum_{j=1}^{n} a_{ij}x_j, \qquad i = 1, 2, \ldots, m,$$

und das Funktional $f(x)$ in der Form

$$f(x) = \sum_{j=1}^{n} f_j x_j$$

schreiben. Aus der Gleichheit

$$f(x) = g(Ax) = \sum_{i=1}^{m} g_i y_i = \sum_{i=1}^{m} \sum_{j=1}^{n} g_i a_{ij} x_j = \sum_{j=1}^{n} x_j \sum_{i=1}^{m} g_i a_{ij}$$

erhalten wir dann $f_j = \sum_{i=1}^{m} g_i a_{ij}$. Wegen $f = A^*g$ folgt hieraus, daß dem Operator A^* die zur Matrix von A transponierte Matrix entspricht.

Die folgenden Eigenschaften des adjungierten Operators ergeben sich unmittelbar aus der Definition:

1. Der Operator A^* ist linear.
2. $(A + B)^* = A^* + B^*$.
3. $(kA)^* = kA^*$ für jede Zahl k.

Ist A ein stetiger Operator aus E in E_1, so ist A^* ein stetiger Operator aus (E_1^*, b) in (E^*, b). (Man beweise das!) Sind E und E_1 Banachräume, so kann diese Aussage folgendermaßen präzisiert werden.

Satz 6. *Ist A ein beschränkter linearer Operator eines Banachraumes E in einen Banachraum E_1, so gilt*

$$\|A^*\| = \|A\|.$$

Beweis. Auf Grund der Eigenschaften der Norm eines Operators gilt

$$|(A^*g, x)| = |(g, Ax)| \leqq \|g\| \cdot \|A\| \cdot \|x\|.$$

Daraus folgt $\|A^*g\| \leqq \|A\| \cdot \|g\|$ bzw.

$$\|A^*\| \leqq \|A\|. \tag{12}$$

Jetzt sei $x \in E$ und $Ax \neq 0$. Setzt man $y_0 = \dfrac{Ax}{\|Ax\|} \in E_1$, so ist offenbar $\|y_0\| = 1$. Nach der Folgerung aus dem Satz von HAHN-BANACH existiert ein Funktional g mit $\|g\| = 1$ und $(g, y_0) = 1$, d. h. $(g, Ax) = \|Ax\|$. Aus der Beziehung

$$\begin{aligned}\|Ax\| &= (g, Ax) = |(A^*g, x)| \\ &\leqq \|A^*g\| \cdot \|x\| \leqq \|A^*\| \cdot \|g\| \cdot \|x\| = \|A^*\| \cdot \|x\|\end{aligned}$$

erhält man dann $\|A\| \leqq \|A^*\|$. Zusammen mit der Ungleichung (12) ergibt sich daraus

$$\|A^*\| = \|A\|,$$

was zu zeigen war.

Aufgabe. E und E_1 seien reflexive Banachräume und $A \in \mathscr{L}(E, E_1)$. Man zeige, daß $A^{**} = A$ ist.

4.5.6. Adjungierte Operatoren im Hilbertraum. Selbstadjungierte Operatoren. Wir wollen den Fall untersuchen, daß A ein beschränkter linearer Operator in einem (reellen oder komplexen) Hilbertraum ist. Nach dem Satz über die allgemeine Form eines linearen Funktionals im Hilbertraum ist die Abbildung τ, die jedem $y \in H$ das lineare Funktional

$$(\tau y)(x) = (x, y)$$

zuordnet, ein Isomorphismus (oder antilinearer Isomorphismus, wenn H komplexer Raum ist) von H auf den dualen Raum H^*. A^* sei der zu A adjungierte Operator. Dann ist klar, daß die Abbildung $\tilde{A}^* = \tau^{-1}A^*\tau$ einen beschränkten linearen Operator in H darstellt. Auch sieht man leicht, daß

$$(Ax, y) = (x, \tilde{A}^*y)$$

für jedes $y \in H$ gilt. Da $\|A\| = \|A^*\|$ ist und τ und τ^{-1} isometrische Abbildungen sind, gilt $\|\tilde{A}^*\| = \|A\|$.

Alles bisher Gesagte ist selbstverständlich auch für einen endlichdimensionalen, reellen oder komplexen, euklidischen Raum gültig.

Wir wollen folgende Vereinbarung treffen. Ist R ein Hilbertraum (oder ein endlichdimensionaler euklidischer Raum), so bezeichnen wir den oben definierten Operator $\tilde{A}^*$ als *adjungierten* Operator von A. Dabei ist also $\tilde{A}^*$ wie A ein Operator, der R in sich abbildet.

Man muß betonen, daß sich diese Definition von der Definition des adjungierten Operators im Fall eines beliebigen Banachraumes E unterscheidet, denn danach wirkt der adjungierte Operator A^* im dualen Raum E^*. Manchmal nennt man den Operator $\tilde{A}^*$ im Unterschied zu A^* auch *hermitesch-adjungiert*. Um jedoch die Terminologie und die Bezeichnungsweise nicht zu komplizieren, werden wir einfach A^* anstelle von $\tilde{A}^*$ schreiben und diesen Operator adjungierten Operator nennen. Dabei werden wir natürlich berücksichtigen, daß der adjungierte Operator im Hilbertraum (bzw. im euklidischen Raum) stets nur in diesem Sinne zu verstehen ist.

Offenbar kann man den zu A adjungierten Operator A^* jetzt durch die Gleichung

$$(Ax, y) = (x, A^*y) \quad \text{für alle } x, y \in R$$

definieren. Da die Operatoren A und A^* in demselben Raum wirken, ist die Gleichheit $A = A^*$ möglich. Durch diese Eigenschaft wird eine wichtige Klasse von Operatoren im Hilbertraum (bzw. im euklidischen Raum) charakterisiert.

Definition 4. Ein beschränkter linearer Operator in einem Hilbertraum (euklidischen Raum) R heißt *selbstadjungiert*, wenn $A = A^*$ gilt, d. h., wenn

$$(Ax, y) = (x, Ay)$$

für alle $x, y \in R$ ist.

Wir erwähnen die folgende wichtige Eigenschaft des zu A adjungierten Operators A^*. Ein Teilraum R_1 eines Hilbertraumes (euklidischen Raumes) R heißt *invariant* bezüglich des Operators A, wenn für alle $x \in R_1$ auch $Ax \in R_1$ ist. *Ist nun der Teilraum R_1 invariant bezüglich A, so ist sein orthogonales Komplement $R_1^\perp$ invariant bezüglich A^*.* Denn für jedes $y \in R_1^\perp$ und alle $x \in R_1$ gilt

$$(x, A^*y) = (Ax, y) = 0$$

wegen $Ax \in R_1$. Für einen selbstadjungierten Operator A ist also das orthogonale Komplement eines beliebigen invarianten Teilraumes ebenfalls ein invarianter Teilraum bezüglich A.

Aufgabe. Man beweise, daß für beschränkte lineare Operatoren A und B in einem Hilbertraum (euklidischen Raum) folgende Gleichheiten gelten:

$$(\alpha A + \beta B)^* = \bar{\alpha} A^* + \bar{\beta} B^*,$$

$$(AB)^* = B^* A^*,$$

$$(A^*)^* = A,$$

$$I^* = I \quad (I \text{ ist der identische Operator}).$$

4.5.7. Das Spektrum eines Operators. Die Resolvente.[1]) In der Operatorentheorie kann man schwerlich einen wichtigeren Begriff als den Begriff des Spektrums angeben. Zunächst wollen wir diesen Begriff im endlichdimensionalen Fall betrachten.

Es sei A ein linearer Operator im n-dimensionalen Raum $\boldsymbol{C}^n$. Eine Zahl λ heißt *Eigenwert* des Operators A, wenn die Gleichung

$$Ax = \lambda x$$

eine von Null verschiedene Lösung besitzt. Die Gesamtheit aller Eigenwerte heißt *Spektrum* des Operators A. Die verbleibenden Werte λ, für die der Operator $A - \lambda I$ also umkehrbar ist, nennt man *regulär*. Der Umkehroperator $(A - \lambda I)^{-1}$ ist dabei auf ganz $\boldsymbol{C}^n$ definiert und, wie jeder Operator in einem endlichdimensionalen Raum, beschränkt. Somit existieren im endlichdimensionalen Raum zwei Möglichkeiten:

1. Die Gleichung $Ax = \lambda x$ besitzt eine von Null verschiedene Lösung, d. h., λ ist ein Eigenwert von A. Dabei existiert der Operator $(A - \lambda I)^{-1}$ nicht.

2. Der Umkehroperator $(A - \lambda I)^{-1}$ existiert, er ist auf dem ganzen Raum definiert und beschränkt, d. h., λ ist ein regulärer Punkt.

Wenn jedoch A ein Operator in einem unendlichdimensionalen Raum E ist, gibt es noch eine dritte Möglichkeit, und zwar die folgende:

3. Der Operator $(A - \lambda I)^{-1}$ existiert, d. h., die Gleichung $Ax = \lambda x$ besitzt nur die triviale Lösung, aber dieser Operator ist nicht auf ganz E erklärt (und möglicherweise unbeschränkt).

Wir führen folgende Terminologie ein. Ist A ein Operator in einem (komplexen) Banachraum, so nennen wir die Zahl λ *regulär* für A, wenn der Operator $R_\lambda = (A - \lambda I)^{-1}$, die sogenannte *Resolvente* von A, auf ganz E definiert und folglich (Satz 3) auch beschränkt ist. Die Gesamtheit der verbleibenden Werte λ heißt *Spektrum* von A. Zum Spektrum gehören alle Eigenwerte von A, denn wenn $(A - \lambda I)\, x = 0$ ist für ein $x \neq 0$, existiert $(A - \lambda I)^{-1}$ nicht. Die Gesamtheit aller Eigenwerte heißt *Punktspektrum*. Der verbleibende Teil des Spektrums, d. h. die Gesamtheit aller λ, für die $(A - \lambda I)^{-1}$ existiert, aber nicht auf ganz E definiert ist, heißt *stetiges Spektrum*. Somit ist jede komplexe Zahl λ für den Operator A entweder regulärer Wert oder Eigenwert oder ein Punkt des stetigen Spektrums. Die Möglichkeit, daß ein Operator noch stetiges Spektrum besitzen kann, ist der wesentliche Unterschied der Operatorentheorie im unendlichdimensionalen Raum gegenüber dem endlichdimensionalen Fall.

Es sei A ein beschränkter Operator in einem Banachraum E. Der Punkt λ sei regulär, d. h. der Operator $(A - \lambda I)^{-1}$ auf ganz E definiert und beschränkt. Dann ist für hinreichend kleines δ auch der Operator $\big(A - (\lambda + \delta)\, I\big)^{-1}$ auf ganz E definiert und beschränkt (Satz 4), d. h., der Punkt $\lambda + \delta$ ist ebenfalls regulär. *Somit bilden*

[1]) Überall dort, wo vom Spektrum eines Operators die Rede ist, nehmen wir an, daß der Operator in einem komplexen Raum wirkt.

die regulären Punkte eine offene Menge. Folglich ist das *Spektrum*, d. h. das Komplement dieser Menge, eine *abgeschlossene Menge.*

Satz 7. *Ist A ein beschränkter linearer Operator im Banachraum E und $|\lambda| > \|A\|$, so ist λ ein regulärer Punkt.*

Beweis. Da offenbar

$$A - \lambda I = -\lambda \left(I - \frac{1}{\lambda} A\right)$$

ist, ergibt sich

$$R_\lambda = (A - \lambda I)^{-1} = -\frac{1}{\lambda}\left(I - \frac{A}{\lambda}\right)^{-1} = -\frac{1}{\lambda} \sum_{k=0}^{\infty} \frac{A^k}{\lambda^k}.$$

Für $\|A\| < |\lambda|$ konvergiert diese Reihe und stellt einen auf ganz E definierten beschränkten Operator dar (Satz 5). *Das Spektrum des Operators A liegt also in einem Kreis mit dem Radius $\|A\|$ um den Nullpunkt.*

Beispiele

1. Im Raum $C[a, b]$ betrachten wir den Operator A, der durch die Formel

$$Ax(t) = tx(t) \tag{13}$$

definiert ist. Dann gilt

$$(A - \lambda I)\, x(t) = (t - \lambda)\, x(t).$$

Da aus der Gleichung

$$(t - \lambda)\, x(t) = 0$$

folgt, daß die stetige Funktion $x(t)$ identisch verschwindet, ist der Operator (13) für jedes λ umkehrbar. Für jedes $\lambda \in [a, b]$ ist jedoch der Umkehroperator, der durch die Formel

$$(A - \lambda I)^{-1}\, x(t) = \frac{1}{t - \lambda}\, x(t)$$

definiert wird, unbeschränkt und nicht auf ganz $C[a,b]$ erklärt. (Man beweise das!) Somit stellt das Intervall $[a, b]$ das Spektrum des Operators (13) dar; dabei existieren keine Eigenwerte, d. h., es ist nur stetiges Spektrum vorhanden.

2. Im Raum l_2 betrachten wir den Operator A, der folgendermaßen definiert ist:

$$A\colon (x_1, x_2, \ldots) \to (0, x_1, x_2, \ldots). \tag{14}$$

Dieser Operator besitzt keinen Eigenwert. (Man beweise das!) Der Operator A^{-1} ist beschränkt, aber nur auf dem Teilraum $x_1 = 0$ von l_2 definiert. Das bedeutet, daß der Punkt $\lambda = 0$ zum Spektrum des Operators gehört.

Aufgabe. Enthält das Spektrum des Operators (14) außer $\lambda = 0$ noch andere Punkte?

Bemerkungen. (1) Jeder beschränkte lineare Operator in einem komplexen Banachraum, der wenigstens ein von Null verschiedenes Element enthält, besitzt ein nichtleeres Spektrum. Es existieren aber Operatoren, deren Spektrum aus einem einzigen Punkt besteht (zum Beispiel der Multiplikationsoperator (13)).

(2) Satz 7 kann folgendermaßen verschärft werden. Es sei

$$r = \lim_{n \to \infty} \sqrt[n]{\|A^n\|}$$

(man kann beweisen, daß dieser Grenzwert für jeden beschränkten Operator A existiert). *Dann liegt das Spektrum von A ganz im Innern des Kreises mit dem Radius r um den Nullpunkt.* Die Zahl r heißt *Spektralradius* von A.

(3) Die Resolventen R_μ und R_λ, die den Punkten μ und λ entsprechen, sind miteinander vertauschbar und genügen der Beziehung

$$R_\mu - R_\lambda = (\mu - \lambda) R_\mu R_\lambda .$$

Die Vertauschbarkeit ergibt sich unmittelbar aus der obigen Gleichheit, die man leicht nachprüfen kann, wenn man beide Seiten mit $(A - \lambda I)(A - \mu I)$ multipliziert. Hieraus folgt, daß in jedem regulären Punkt λ_0 von A der Grenzwert

$$\lim_{\Delta\lambda \to 0} \frac{R_{\lambda_0 + \Delta\lambda} - R_{\lambda_0}}{\Delta\lambda}$$

(im Sinne der Operatornorm) existiert und gleich $R_{\lambda_0}^2$ ist. Diesen Grenzwert bezeichnet man auch als *Ableitung* von R_λ nach λ.

Aufgabe. A sei ein beschränkter selbstadjungierter Operator in einem komplexen Hilbertraum H. Man beweise, daß das Spektrum von A eine beschränkte abgeschlossene Teilmenge der reellen Achse ist.

4.6. Kompakte Operatoren

4.6.1. Definition und Beispiele kompakter Operatoren. Lineare Operatoren in endlichdimensionalen Räumen kann man vollständig beschreiben. Im Gegensatz dazu stellt die Untersuchung beliebiger linearer Operatoren in unendlichdimensionalen Räumen ein äußerst schwieriges und im wesentlichen unüberschaubares Problem dar. Jedoch können einige wichtige Klassen solcher Operatoren vollständig beschrieben werden. Eines der wichtigsten Beispiele bilden dabei die sogenannten kompakten Operatoren. Diese Operatoren stehen in ihren Eigenschaften einerseits den *ausgearteten* Operatoren (d. h. den beschränkten Operatoren mit endlichdimensionalem Bildraum) sehr nahe und gestatten eine hinreichend detaillierte Beschreibung. Andererseits spielen sie in verschiedenen Anwendungen eine wichtige Rolle, und zwar in erster Linie in der Theorie der Integralgleichungen, denen Kapitel 9 gewidmet ist.

Definition 1. Ein Operator A, der einen Banachraum E in sich (oder in einen Banachraum E_1) abbildet, heißt *kompakt* oder *vollstetig*, wenn er jede beschränkte Menge in eine präkompakte Menge überführt.

In einem endlichdimensionalen normierten Raum ist jeder lineare Operator kompakt. Denn er führt jede beschränkte Menge wieder in eine beschränkte Menge über, und jede solche Menge ist in einem endlichdimensionalen Raum präkompakt.

In einem unendlichdimensionalen Raum ist die Kompaktheit eines Operators eine wesentlich stärkere Forderung als nur die Stetigkeit (d. h. die Beschränktheit). Beispielsweise ist der identische Operator in einem Hilbertraum stetig, aber niemals kompakt. (Man beweise das unabhängig vom unten betrachteten Beispiel 1.) Wir wollen einige Beispiele betrachten.

1. I sei der identische Operator des Banachraumes E. Wir werden zeigen, daß der Operator I nicht kompakt ist, falls E unendlichdimensional ist. Dazu braucht man offenbar nur zu zeigen, daß die Einheitskugel in E (die natürlich von I in sich übergeführt wird) nicht kompakt ist. Das wiederum ergibt sich aus dem folgenden Lemma, das wir im weiteren noch benötigen.

Lemma 1. *Es seien $x_1, x_2, \dots$ linear unabhängige Vektoren in einem normierten Raum E, und es sei E_n der durch die Vektoren $x_1, \dots, x_n$ erzeugte Teilraum. Dann existiert eine Folge $y_1, y_2, \dots$ von Vektoren mit folgenden Eigenschaften:*

$$\text{a) } \|y_n\| = 1; \text{ b) } y_n \in E_n; \text{ c) } \varrho(y_n, E_{n-1}) > \frac{1}{2}.$$

Dabei ist $\varrho(y_n, E_{n-1})$ der Abstand des Vektors y_n von E_{n-1}, d. h.

$$\inf_{x \in E_{n-1}} \|y_n - x\|.$$

Beweis. Da die Vektoren $x_1, x_2, \dots$ linear unabhängig sind, ist $x_n \notin E_{n-1}$ und $\varrho(x_n, E_{n-1}) = \alpha > 0$. Es sei nun x^* ein Vektor aus E_{n-1} mit $\|x_n - x^*\| < 2\alpha$. Wegen $\alpha = \varrho(x_n, E_{n-1}) = \varrho(x_n - x^*, E_{n-1})$ erfüllt der Vektor

$$y_n = \frac{x_n - x^*}{\|x_n - x^*\|}$$

dann die Bedingungen a) bis c). Als y_1 kann man dabei $\dfrac{x_1}{\|x_1\|}$ nehmen. Damit ist das Lemma bewiesen.

Wenn man dieses Lemma auf die Einheitskugel eines unendlichdimensionalen normierten Raumes anwendet, kann man also eine Folge $\{y_n\}$ konstruieren mit $\varrho(y_{n-1}, y_n) > 1/2$. Eine solche Folge enthält offenbar keine konvergente Teilfolge. Das bedeutet aber, daß die Einheitskugel nicht kompakt ist.

2. A sei ein stetiger linearer Operator eines Banachraumes E in einen endlichdimensionalen Teilraum von E. Ein solcher Operator ist kompakt, da er jede beschränkte Menge $M \subset E$ wieder in eine beschränkte Menge des endlichdimensionalen Teilraumes überführt. Dort ist aber jede beschränkte Menge präkompakt.

Insbesondere ist die orthogonale Projektion eines Hilbertraumes auf einen Teilraum genau dann kompakt, wenn dieser Teilraum endlichdimensional ist.

3. Wir betrachten im Raum l_2 den durch die Formel

$$Ax = \left(x_1, \frac{1}{2}\,x_2, \ldots, \frac{1}{2^n}\,x_n, \ldots\right) \tag{1}$$

für jedes $x = (x_1, x_2, \ldots, x_n, \ldots)$ definierten Operator A. Dieser Operator ist kompakt. Da jede beschränkte Menge aus l_2 in einer gewissen Kugel dieses Raumes enthalten ist, brauchen wir nur zu zeigen, daß die Bilder von Kugeln präkompakt sind. Wegen der Linearität von A genügt es, das für die Einheitskugel nachzuweisen. Der Operator (1) führt nun aber die Einheitskugel des Raumes l_2 in eine Teilmenge des Fundamentalquaders über (siehe 2.7.1.). Folglich ist diese Menge totalbeschränkt und damit präkompakt.

Aufgabe. Es sei $Ax = (a_1x_1, a_2x_2, \ldots, a_nx_n, \ldots)$. Unter welchen Bedingungen an die Zahlenfolge $\{a_n\}$ ist dieser Operator kompakt in l_2?

4. Eine wichtige Klasse kompakter Operatoren im Raum $C[a, b]$ der stetigen Funktionen wird von Operatoren gebildet, die darstellbar sind in der Form

$$Ax = y(s) = \int_a^b K(s, t)\, x(t)\, dt. \tag{2}$$

Wir wollen die Gültigkeit der folgenden Aussage zeigen: *Ist die Funktion $K(s, t)$ im Quadrat $a \leqq s \leqq b$, $a \leqq t \leqq b$ beschränkt und liegen alle ihre Unstetigkeitsstellen auf endlich vielen Kurven*

$$t = \varphi_k(s), \qquad k = 1, 2, \ldots, n,$$

wobei φ_k stetige Funktionen sind, so definiert Formel (2) *einen kompakten Operator im Raum $C[a, b]$.*

Dazu vermerken wir erst einmal, daß das Integral (2) unter den angegebenen Bedingungen für jedes s aus dem Intervall $[a, b]$ existiert, d. h., daß die Funktion $y(s)$ sinnvoll definiert ist. Es seien nun

$$M = \sup_{a \leqq s, t \leqq b} |K(s, t)|$$

und G die Menge aller der Punkte (s, t), die für wenigstens ein $k = 1, 2, \ldots, n$ der Ungleichung

$$|t - \varphi_k(s)| < \frac{\varepsilon}{12Mn}$$

genügen. Die Spur $G(s)$ dieser Menge auf der Geraden $s = \text{const}$ läßt sich als Vereinigung

$$G(s) = \bigcup_{k=1}^{n} \left\{t\colon |t - \varphi_k(s)| < \frac{\varepsilon}{12Mn}\right\}$$

von Intervallen darstellen. Ferner sei F das Komplement von G bezüglich des Quadrates $a \leqq s, t \leqq b$. Wegen der Stetigkeit von $K(s, t)$ auf der kompaktenMen ge F existiert ein $\delta > 0$, so daß

$$|K(s', t') - K(s'', t'')| < \frac{\varepsilon}{3(b-a)}$$

ist für alle Punkte (s', t'), (s'', t'') aus F mit

$$|s' - s''| + |t' - t''| < \delta. \tag{3}$$

Wir schätzen jetzt die Differenz $y(s') - y(s'')$ unter der Voraussetzung $|s' - s''| < \delta$ ab. Zunächst ist

$$|y(s') - y(s'')| \leqq \int_a^b |K(s', t) - K(s'', t)| \cdot |x(t)| \, dt.$$

Zur Abschätzung des rechts stehenden Integrals zerlegen wir den Integrationsbereich $[a, b]$ in die Mengen $P = G(s') \cup G(s'')$ und $Q = [a, b] \setminus P$. Da P eine Vereinigung von Intervallen ist, deren Gesamtlänge $\frac{\varepsilon}{3M}$ nicht übersteigt, erhalten wir

$$\int_P |K(s', t) - K(s'', t)| \cdot |x(t)| \, dt < \frac{2\varepsilon}{3} \|x\|.$$

Das Integral über Q gestattet offenbar die Abschätzung

$$\int_Q |K(s', t) - K(s'', t)| \cdot |x(t)| \, dt < \frac{\varepsilon}{3} \|x\|.$$

Somit gilt

$$|y(s') - y(s'')| < \varepsilon \|x\|. \tag{4}$$

Ungleichung (4) zeigt, daß die Funktion $y(s)$ stetig ist. Formel (2) definiert also einen Operator, der $C[a, b]$ in sich überführt. Ferner ist aus dieser Ungleichung ersichtlich, daß für jede in $C[a, b]$ beschränkte Menge $\{x(t)\}$ die entsprechende Menge $\{y(s)\}$ gleichgradig stetig ist. Schließlich folgt aus $\|x\| \leqq C$

$$\|y\| = \sup |y(s)| \leqq \sup \int_a^b |K(s, t)| \cdot |x(t)| \, dt \leqq M(b-a) \|x\|.$$

Somit führt der Operator (2) jede beschränkte Menge aus $C[a, b]$ in eine Menge gleichmäßig beschränkter und gleichgradig stetiger Funktionen, d. h. in eine präkompakte Menge über.

4a. Die Lage der Unstetigkeitspunkte der Funktion $K(s, t)$ auf endlich vielen Kurven, die die Geraden $s = \text{const}$ jeweils nur in einem Punkt schneiden, ist wesent-

lich. Es sei zum Beispiel

$$K(s, t) = \begin{cases} 1 & \text{für} \quad s < 1/2, \\ 0 & \text{für} \quad s \geqq 1/2 \end{cases}$$

im Quadrat $0 \leqq s, t \leqq 1$. Der Operator (2) mit diesem Kern, der auf der ganzen Strecke $s = 1/2$, $0 \leqq t \leqq 1$, Unstetigkeitsstellen besitzt, führt die Funktion $x(t) \equiv 1$ in eine unstetige Funktion über.

4b. Setzt man $K(s, t) = 0$ für $t < s$, so nimmt der Operator (2) die Gestalt

$$y(s) = \int_a^s K(s, t)\, x(t)\, dt \tag{5}$$

an. Wenn wir annehmen, daß die Funktion $K(s, t)$ für $t < s$ stetig ist, folgt nach dem in Beispiel 4 Gesagten, daß der Operator (2) in $C[a, b]$ vollstetig ist. Dieser Operator heißt *Operator vom Volterraschen*[1]) *Typ*.

Bemerkung. Bei der von uns benutzten Definition eines kompakten Operators kann es vorkommen, daß das Bild der abgeschlossenen Einheitskugel nicht kompakt ist (obwohl es präkompakt ist). Wir betrachten zum Beispiel im Raum $C[-1, 1]$ den Integrationsoperator

$$Ix(s) = \int_{-1}^s x(t)\, dt.$$

Nach dem oben Bewiesenen ist I ein vollstetiger Operator in $C[-1, 1]$. Wenn wir

$$x_n(t) = \begin{cases} 0 & \text{für} \quad -1 \leqq t \leqq 0, \\ nt & \text{für} \quad 0 < t \leqq \dfrac{1}{n}, \\ 1 & \text{für} \quad \dfrac{1}{n} < t \leqq 1 \end{cases}$$

setzen, so ist $x_n \in C[-1, 1]$, $\|x_n\| = 1$ für alle n und

$$y_n(t) = Ix_n(t) = \begin{cases} 0 & \text{für} \quad -1 \leqq t \leqq 0, \\ \dfrac{1}{2} nt^2 & \text{für} \quad 0 < t \leqq \dfrac{1}{n}, \\ t - \dfrac{1}{2n} & \text{für} \quad \dfrac{1}{n} < t \leqq 1. \end{cases}$$

Offenbar konvergiert die Folge y_n in $C[-1, 1]$ gegen die Funktion

$$y(t) = \begin{cases} 0 & \text{für} \quad -1 \leqq t \leqq 0, \\ t & \text{für} \quad 0 < t \leqq 1. \end{cases}$$

---

[1]) Vito Volterra (1860—1940), italienischer Mathematiker, Autor einer Reihe von Arbeiten über Funktionalanalysis und Integralgleichungen.

Diese Funktion kommt aber nicht als Bild (bei der Abbildung I) einer Funktion aus $C[-1, 1]$ vor, da die Funktion $y'(t)$ unstetig ist.

Jedoch kann man im Fall eines reflexiven Raumes (z. B. eines Hilbertraumes) beweisen, daß das Bild der abgeschlossenen Einheitskugel bei einer kompakten linearen Abbildung kompakt ist.

4.6.2. Grundlegende Eigenschaften kompakter Operatoren.

Satz 1. *Ist $\{A_n\}$ eine Folge kompakter Operatoren in einem Banachraum E, die in der Norm gegen einen Operator A konvergiert, so ist A ebenfalls kompakt.*

Beweis. Zum Nachweis der Kompaktheit des Operators A genügt es zu zeigen, daß man aus der Bildfolge $\{Ax_n\}$ einer beliebigen beschränkten Folge $\{x_n\}$ aus E stets eine konvergente Teilfolge auswählen kann.

Da der Operator A_1 kompakt ist, kann man aus der Folge $\{x_n\}$ eine Teilfolge

$$x_1^{(1)}, x_2^{(1)}, \ldots, x_n^{(1)}, \ldots \tag{6}$$

auswählen, so daß die Folge $\{A_1 x_n^{(1)}\}$ konvergent ist. Aus der Folge $\{x_n^{(1)}\}$ kann man wiederum eine Teilfolge

$$x_1^{(2)}, x_2^{(2)}, \ldots, x_n^{(2)}, \ldots$$

auswählen, so daß die Folge $\{A_2 x_n^{(2)}\}$ konvergent ist. Dabei konvergiert natürlich auch $\{A_1 x_n^{(2)}\}$. Analog wählen wir aus der Folge $\{x_n^{(2)}\}$ eine Teilfolge

$$x_1^{(3)}, x_2^{(3)}, \ldots, x_n^{(3)}, \ldots$$

aus, so daß $\{A_3 x_n^{(3)}\}$ konvergent ist usw. Anschließend nehmen wir die Diagonalfolge

$$x_1^{(1)}, x_2^{(2)}, \ldots, x_n^{(n)}, \ldots$$

Diese Folge wird von jedem der Operatoren $A_1, A_2, \ldots, A_n, \ldots$ in eine konvergente Folge übergeführt. Wir werden zeigen, daß auch der Operator A diese Folge in eine konvergente Folge überführt. Damit ist dann die Kompaktheit von A nachgewiesen. Da der Raum E vollständig ist, genügt es zu zeigen, daß $\{Ax_n^{(n)}\}$ eine Fundamentalfolge ist. Es gilt

$$\begin{aligned}\|Ax_n^{(n)} - Ax_m^{(m)}\| \leqq \|Ax_n^{(n)} - A_k x_n^{(n)}\| + \|A_k x_n^{(n)} - A_k x_m^{(m)}\| \\ + \|A_k x_m^{(m)} - Ax_m^{(m)}\|.\end{aligned} \tag{7}$$

Nun sei $\|x_n\| \leqq C$. Dann wählen wir zunächst k so, daß $\|A - A_k\| < \dfrac{\varepsilon}{3C}$ ist. Anschließend wählen wir N so, daß für $n, m > N$ die Ungleichung

$$\|A_k x_n^{(n)} - A_k x_m^{(m)}\| < \frac{\varepsilon}{3}$$

erfüllt ist. (Das ist möglich, da die Folge $\{A_k x_n^{(n)}\}$ konvergent ist.) Unter diesen Bedingungen erhalten wir aus (7), daß

$$\|Ax_n^{(n)} - Ax_m^{(m)}\| < \varepsilon$$

ist für alle hinreichend großen n und m. Damit ist der Satz bewiesen.

Man prüft leicht nach, daß Linearkombinationen kompakter Operatoren ebenfalls kompakt sind. Folglich bilden die kompakten Operatoren im Raum $\mathcal{L}(E, E)$ aller beschränkten linearen Operatoren auf E einen abgeschlossenen linearen Teilraum.

Wir wollen jetzt untersuchen, ob die Gesamtheit der kompakten Operatoren auch bezüglich der Nacheinanderausführung von Operatoren abgeschlossen ist. Hier gilt sogar eine wesentlich schärfere Aussage.

Satz 2. *Ist A ein kompakter und B ein beschränkter Operator, so sind die Operatoren AB und BA kompakt.*

Beweis. Die Menge $M \subset E$ sei beschränkt. Dann ist die Menge BM beschränkt und daher die Menge ABM präkompakt. Also ist der Operator AB kompakt. Andererseits ist die Menge AM präkompakt, und auf Grund der Stetigkeit von B ist auch BAM präkompakt, d. h., der Operator BA ist kompakt. Damit ist der Satz bewiesen.

Folgerung. *In einem unendlichdimensionalen Raum E kann ein kompakter Operator keinen beschränkten Umkehroperator besitzen.*

Sonst wäre nämlich der identische Operator $I = A^{-1}A$ kompakt in E, was aber unmöglich ist (siehe Beispiel 1).

Bemerkung. Satz 2 zeigt, daß die kompakten Operatoren im Ring $\mathcal{L}(E, E)$ aller beschränkten Operatoren ein zweiseitiges Ideal[1]) bilden.

Satz 3. *Der adjungierte Operator eines kompakten Operators ist kompakt.*

Beweis. A sei ein kompakter Operator im Banachraum E. Wir wollen zeigen, daß der adjungierte Operator A^* in E^* jede beschränkte Menge in eine präkompakte Menge überführt. Da jede beschränkte Teilmenge eines normierten Raumes in einer hinreichend großen Kugel enthalten ist, genügt der Nachweis, daß A^* jede Kugel in eine präkompakte Menge überführt. Wegen der Linearität des Operators A^* braucht man nur zu zeigen, daß das Bild A^*S^* der abgeschlossenen Einheitskugel $S^* \subset E^*$ präkompakt ist.

Dazu betrachten wir die Elemente aus E^* als Funktionen, und zwar nicht auf dem ganzen Raum E, sondern nur auf dem Kompaktum $\overline{AS}$, dem Abschluß des Bildes der Einheitskugel $S \subset E$ bei der Abbildung A. Dabei ist die Menge Φ der Funktionen, die einem Funktional aus S^* entsprechen, gleichmäßig beschränkt und gleichgradig stetig. Das ergibt sich aus den für $\|\varphi\| \leqq 1$ gültigen Abschätzungen

$$\sup_{x \in \overline{AS}} |\varphi(x)| = \sup_{x \in AS} |\varphi(x)| \leqq \|\varphi\| \sup_{x \in S} \|Ax\| \leqq \|A\|$$

[1]) Ein Unterring $\mathcal{U}$ eines Ringes R heißt (zweiseitiges) *Ideal* in R, wenn mit $a \in \mathcal{U}, r \in R$ stets auch ar und ra zu $\mathcal{U}$ gehören.

und

$$|\varphi(x') - \varphi(x'')| \leqq \|\varphi\| \cdot \|x' - x''\| \leqq \|x' - x''\|.$$

Folglich ist die Menge Φ im Raum $C[\overline{AS}]$ (nach dem Satz von Arzelà) präkompakt. Nun ist aber die Menge Φ mit der Metrik, die durch die gewöhnliche Metrik des Raumes $C[\overline{AS}]$ induziert wird, isometrisch zur Menge A^*S^* (mit der durch die Norm von E^* induzierten Metrik). Denn für $g_1, g_2 \in S^*$ gilt

$$\begin{aligned}\|A^*g_1 - A^*g_2\| &= \sup_{x\in S} |(A^*g_1 - A^*g_2, x)| \\ &= \sup_{x\in S} |(g_1 - g_2, Ax)| \\ &= \sup_{z\in AS} |(g_1 - g_2, z)| \\ &= \sup_{z\in \overline{AS}} |(g_1 - g_2, z)| = \varrho(g_1, g_2).\end{aligned}$$

Wegen der Präkompaktheit ist Φ totalbeschränkt, folglich ist auch die zu Φ isometrische Menge A^*S^* totalbeschränkt. Daher ist A^*S^* präkompakt in E^*. Damit ist der Satz bewiesen.

Bemerkung. Es ist nicht schwer nachzuweisen, daß die Menge Φ in $C[\overline{AS}]$ abgeschlossen ist. Somit ist Φ kompakt, und daher ist auch die Menge A^*S^* kompakt, obwohl (wie aus der Bemerkung auf S. 238 ersichtlich ist) das Bild der abgeschlossenen Einheitskugel bei einer beliebigen vollstetigen Abbildung nicht kompakt zu sein braucht. Die Situation im eben bewiesenen Satz unterscheidet sich vom allgemeinen Fall dadurch, daß die abgeschlossene Einheitskugel S^* in E^* kompakt in der w^*-Topologie des Raumes E^* ist (siehe 4.3., Satz 5). Hieraus folgt aber die Kompaktheit des Bildes der Menge S^* (in der Metrik des Raumes E^*) für jeden kompakten Operator.

Aufgaben

1. A sei ein beschränkter linearer Operator in einem Banachraum. Man beweise, daß aus der Kompaktheit von A^* die Kompaktheit von A folgt.

2. Ein linearer Operator A im Hilbertraum H ist genau dann kompakt, wenn der zu A (hermitesch) adjungierte Operator A^* kompakt ist.

4.6.3. Eigenwerte eines kompakten Operators.

Satz 4. *A sei ein kompakter Operator in einem Banachraum E. Dann gibt es zu jedem $\delta > 0$ nur endlich viele linear unabhängige Eigenvektoren, so daß für die zugehörigen Eigenwerte λ die Beziehung $|\lambda| > \delta$ erfüllt ist.*[1]

Beweis. Es sei $\lambda_1, \lambda_2, \ldots, \lambda_n, \ldots$ eine Folge von (nicht notwendig verschiedenen) Eigenwerten des Operators A mit $|\lambda_n| > \delta$. Ferner sei $x_1, x_2, \ldots, x_n, \ldots$ eine zu-

[1]) Anm. d. Übers.: Eine (reelle oder komplexe) Zahl λ heißt *Eigenwert*, falls es einen Vektor $x = 0$ mit $Ax = \lambda x$ gibt. Derartige Vektoren heißen *Eigenvektoren*.

gehörige Folge von Eigenvektoren, $Ax_n = \lambda_n x_n$. Diese Vektoren seien linear unabhängig.

Nach Lemma 1 von S. 235 gibt es dann eine Folge von Vektoren $y_1, y_2, \ldots, y_n, \ldots$ mit

$$\text{a) } y_n \in E_n; \quad \text{b) } \|y_n\| = 1; \quad \text{c) } \varrho(y_n, E_n) = \inf_{x \in E_{n-1}} \|y_n - x\| > \frac{1}{2},$$

wobei E_n der von den Vektoren $x_1, x_2, \ldots, x_n$ erzeugte Teilraum ist.

Wegen der Ungleichung $|\lambda_n| > \delta$ ist die Folge $\left\{\frac{y_n}{\lambda_n}\right\}$ beschränkt. Aus der Folge $\left\{A\left(\frac{y_n}{\lambda_n}\right)\right\}$ der zugehörigen Bilder kann man jedoch keine konvergente Teilfolge auswählen. Da nämlich $y_n = \sum_{k=1}^{n} \alpha_k x_k$ ist, gilt

$$A\left(\frac{y_n}{\lambda_n}\right) = \sum_{k=1}^{n-1} \frac{\alpha_k \lambda_k}{\lambda_n} x_k + \alpha_n x_n = y_n + z_n$$

mit

$$z_n = \sum_{k=1}^{n-1} \alpha_k \left(\frac{\lambda_k}{\lambda_n} - 1\right) x_k \in E_{n-1},$$

und daher folgt für beliebige $p > q$

$$\left\|A\left(\frac{y_p}{\lambda_p}\right) - A\left(\frac{y_q}{\lambda_q}\right)\right\| = \|y_p + z_p - (y_q + z_q)\|$$

$$= \|y_p - (y_q + z_q - z_p)\| > \frac{1}{2},$$

weil $y_q + z_q - z_p$ zu E_{p-1} gehört. Das widerspricht aber der Kompaktheit des Operators A.

Aus dem bewiesenen Satz ergibt sich insbesondere, daß die Anzahl der linear unabhängigen Eigenvektoren, die zu einem gegebenen Eigenwert $\lambda \neq 0$ gehören, stets endlich ist.

Weiterhin folgt aus diesem Satz, daß ein kompakter Operator A für jede positive Zahl δ jeweils nur endlich viele Eigenwerte λ_n mit $|\lambda_n| > \delta$ besitzt, und daß man alle die Eigenwerte von A nach abnehmendem Betrag in einer Folge anordnen kann: $|\lambda_1| \geqq |\lambda_2| \geqq \cdots$.

4.6.4. Kompakte Operatoren im Hilbertraum. Bisher haben wir kompakte Operatoren in einem beliebigen Banachraum betrachtet. Wir wollen diese Betrachtungen jetzt durch einige Fakten vervollständigen, die sich auf kompakte Operatoren in einem Hilbertraum beziehen.

Wir haben einen Operator kompakt genannt, wenn er jede beschränkte Menge in eine präkompakte Menge überführt. Jeder (separable) Hilbertraum ist wegen $H = H^*$ der duale Raum eines separablen Raumes, und daher sind alle beschränkten Mengen in H (und nur diese) schwach kompakt. Somit kann man einen kompakten Operator im Hilbertraum auch als einen Operator definieren, der jede schwachkompakte Menge in eine (bezüglich der starken Topologie) präkompakte Menge überführt.

Schließlich ist in einigen Fällen auch noch die folgende Definition der Kompaktheit eines Operators zweckmäßig: Ein Operator A heißt *kompakt* in H, wenn er jede schwachkonvergente Folge in eine starkkonvergente überführt.

Wir wollen die Gleichwertigkeit beider Definitionen nachweisen. Zunächst sei die letztere Bedingung erfüllt, und M sei eine beschränkte Menge in H. Jede unendliche Teilmenge von M enthält dann eine schwachkonvergente Folge. Diese Folge wird durch A in eine starkkonvergente Folge übergeführt, d. h., AM ist präkompakt. Umgekehrt seien A ein kompakter Operator, $\{x_n\}$ eine schwachkonvergente Folge und x deren schwacher Grenzwert. Dann enthält $\{AX_n\}$ eine starkkonvergente Teilfolge. Gleichzeitig ist $\{Ax_n\}$ wegen der Stetigkeit von A schwach konvergent gegen Ax, so daß $\{Ax_n\}$ nicht mehr als einen Häufungspunkt besitzen kann. Folglich ist $\{Ax_n\}$ eine konvergente Folge.

4.6.5. Selbstadjungierte Operatoren im Hilbertraum. Für selbstadjungierte Operatoren in einem endlichdimensionalen euklidischen Raum ist bekannt, daß man stets eine geeignete orthonormierte Basis finden kann, so daß die dem Operator bezüglich dieser Basis zugeordnete Matrix eine Diagonalmatrix ist. In diesem Abschnitt wollen wir dieses Resultat auf kompakte selbstadjungierte Operatoren in einem Hilbertraum ausdehnen. Die Ergebnisse dieses Abschnitts sind sowohl für den reellen als auch für den komplexen Fall gültig.

Zunächst weisen wir auf einige Eigenschaften von Eigenvektoren und Eigenwerten selbstadjungierter Operatoren in einem Hilbertraum hin, die übrigens den entsprechenden Eigenschaften selbstadjungierter Operatoren im endlichdimensionalen Fall vollkommen analog sind.

I. *Alle Eigenwerte eines selbstadjungierten Operators A in H sind reell.*

Ist nämlich $Ax = \lambda x$, $\|x\| \neq 0$, so gilt

$$\lambda(x, x) = (Ax, x) = (x, Ax) = (x, \lambda x) = \bar{\lambda}(x, x),$$

woraus $\lambda = \bar{\lambda}$ folgt.

II. *Eigenvektoren zu verschiedenen Eigenwerten sind zueinander orthogonal.*

Aus $Ax = \lambda x$ und $Ay = \mu y$ folgt nämlich

$$\lambda(x, y) = (Ax, y) = (x, Ay) = (x, \mu y) = \mu(x, y),$$

woraus sich $(x, y) = 0$ für $\lambda \neq \mu$ ergibt.

Jetzt werden wir den folgenden fundamentalen Satz beweisen.

Satz 5 (Hilbert-Schmidt). *Für jeden kompakten selbstadjungierten linearen Operator A im Hilbertraum H existiert ein orthonormiertes System $\{\varphi_n\}$ von Eigenvektoren, so daß jedes Element $\xi \in H$ eindeutig in der Form*

$$\xi = \sum_k c_k \varphi_k + \xi'$$

mit einem Vektor $\xi' \in \operatorname{Ker} A$ (d. h. $A\xi' = 0$) dargestellt werden kann. Dabei gilt

$$A\xi = \sum_k \lambda_k c_k \varphi_k,$$

wobei λ_k der φ_k entsprechende Eigenwert ist. Ist das System $\{\varphi_k\}$ unendlich, so gilt

$$\lim_{n\to\infty} \lambda_n = 0.$$

Zum Beweis dieses grundlegenden Satzes benötigen wir folgende Hilfssätze.

Lemma 2. *Ist $\{\xi_n\}$ schwach konvergent gegen ξ und ist der lineare Operator A kompakt, so gilt*

$$Q(\xi_n) = (A\xi_n, \xi_n) \to (A\xi, \xi) = Q(\xi).$$

Beweis. Für jedes n gilt

$$|(A\xi_n, \xi_n) - (A\xi, \xi)| \leqq |(A\xi_n, \xi_n) - (A\xi, \xi_n)| + |(A\xi, \xi_n) - (A\xi, \xi)|.$$

Dabei ist

$$|(A\xi_n, \xi_n) - (A\xi, \xi_n)| \leqq \|\xi_n\| \cdot \|A(\xi_n - \xi)\|$$

und

$$|(A\xi_n, \xi) - (A\xi, \xi)| = |(\xi, A(\xi_n - \xi)| \leqq \|\xi\| \cdot \|A(\xi_n - \xi)\|.$$

Da hier die Zahlen $\|\xi_n\|$ beschränkt sind, während $\|A(\xi_n - \xi)\| \to 0$ strebt, ergibt sich

$$|(A\xi_n, \xi_n) - (A\xi, \xi)| \to 0,$$

was zu zeigen war.

Lemma 3. *A sei ein beschränkter selbstadjungierter linearer Operator. Wenn das Funktional*

$$|Q(\xi)| = |(A\xi, \xi)|$$

sein Maximum auf der Einheitskugel im Punkt ξ_0 annimmt, folgt aus

$$(\xi_0, \eta) = 0$$

die Beziehung

$$(A\xi_0, \eta) = (\xi_0, A\eta) = 0.$$

Beweis. Wir setzen

$$\xi = \frac{\xi_0 + a\eta}{\sqrt{1 + |a|^2 \, \|\eta\|^2}}$$

mit einer beliebigen komplexen Zahl a. Nach Voraussetzung ist $(\xi_0, \eta) = 0$ und $\|\xi_0\| = 1$, daher ist auch $\|\xi\| = 1$. Wegen

$$Q(\xi) = \frac{1}{1 + |a|^2 \, \|\eta\|^2} \, [Q(\xi_0) + 2a(A\xi_0, \eta) + a^2 Q(\eta)]$$

gilt für hinreichend kleine a

$$Q(\xi) = Q(\xi_0) + 2a(A\xi_0, \eta) + O(a^2).$$

Auf Grund dieser Beziehung kann man a im Fall $(A\xi_0, \eta) \neq 0$ offenbar so wählen, daß $|Q(\xi)| > Q(\xi_0)$ ist. Das widerspricht aber der Voraussetzung des Lemmas.

Aus Lemma 3 ergibt sich unmittelbar, daß ξ_0 ein Eigenvektor des Operators ist, wenn $|Q(\xi)|$ sein Maximum für $\xi = \xi_0$ annimmt.

Beweis von Satz 5. Wir werden die Elemente φ_k induktiv konstruieren, so daß die absoluten Beträge der entsprechenden Eigenwerte dabei eine fallende Folge bilden:

$$|\lambda_1| \geqq |\lambda_2| \geqq \cdots \geqq |\lambda_n| \cdots.$$

Zur Konstruktion des Elements φ_1 betrachten wir den Ausdruck $|Q(\xi)| = |(A\xi, \xi)|$. Wir wollen beweisen, daß er sein Maximum auf der Einheitskugel annimmt. Dazu sei

$$S = \sup_{\|\xi\| \leqq 1} |(A\xi, \xi)|$$

und $\xi_1, \xi_2, \ldots$ eine Folge mit $\|\xi_n\| \leqq 1$ und

$$|(A\xi_n, \xi_n)| \to S \quad \text{für} \quad n \to \infty.$$

Da die Einheitskugel in H schwachkompakt ist, kann man aus der Folge $\{\xi_n\}$ eine Teilfolge auswählen, die schwach gegen ein gewisses Element η konvergiert. Dabei ist $\|\eta\| \leqq 1$, und auf Grund von Lemma 2 gilt

$$|(A\eta, \eta)| = S.$$

Hier ist sogar $\|\eta\| = 1$. (Ist nämlich $\|\eta\| < 1$, so können wir $\eta_1 = \dfrac{\eta}{\|\eta\|}$ setzen und erhalten $\|\eta_1\| = 1$ und $|(A\eta_1, \eta_1)| > S$, was der Definition von S widerspricht.) Als φ_1 nehmen wir nun dieses Element η. Dann gilt

$$A\varphi_1 = \lambda_1 \varphi_1,$$

woraus

$$|\lambda_1| = \frac{|(A\varphi_1, \varphi_1)|}{(\varphi_1, \varphi_1)} = |(A\varphi_1, \varphi_1)| = S$$

folgt. Jetzt seien die Eigenvektoren

$$\varphi_1, \varphi_2, \ldots, \varphi_n$$

mit den zugehörigen Eigenwerten

$$\lambda_1, \lambda_2, \ldots, \lambda_n$$

bereits konstruiert. $M(\varphi_1, \varphi_2, \ldots, \varphi_n)$ sei der von den Vektoren $\varphi_1, \varphi_2, \ldots, \varphi_n$ aufgespannte Teilraum. Dann betrachten wir das Funktional

$$|(A\xi, \xi)|$$

auf der Gesamtheit der Elemente ξ aus

$$M_n{}^{\perp} = H \ominus M(\varphi_1, \varphi_2, \ldots, \varphi_n)$$

(die zu den Vektoren $\varphi_1, \varphi_2, \ldots, \varphi_n$ orthogonal sind) mit $\|\xi\| < 1$. Die Menge $M_n{}^{\perp}$ ist ein bezüglich A invarianter Teilraum (da M ein invarianter Teilraum ist und A selbstadjungiert ist). Also können wir die obigen Betrachtungen auf $M_n{}^{\perp}$ anwenden und erhalten in $M_n{}^{\perp}$ einen Eigenvektor von A (den wir φ_{n+1} nennen).

Es sind nun zwei Fälle möglich:

1. Nach endlich vielen Schritten erhalten wir einen Teilraum $M_{n_0}^{\perp}$, in dem $(A\xi, \xi) = 0$ ist;

2. $(A\xi, \xi) \not\equiv 0$ auf $M_n{}^{\perp}$ für alle n.

Im ersten Fall folgt aus Lemma 3, daß der Operator A den Teilraum $M_{n_0}^{\perp}$ in die Null überführt. $M_{n_0}^{\perp}$ besteht also ganz aus Eigenvektoren zum Eigenwert $\lambda = 0$. Das System der Vektoren $\{\varphi_n\}$ besteht dann aus endlich vielen Elementen.

Im zweiten Fall erhalten wir eine Folge $\{\varphi_n\}$ von Eigenvektoren, deren Eigenwerte λ_n alle von Null verschieden sind. Wir zeigen, daß dann $\lambda_n \to 0$ strebt. Die Folge $\{\varphi_n\}$ konvergiert (wie jede orthonormierte Folge) schwach gegen Null. Daher müssen die $A\varphi_n = \lambda_n \varphi_n$ in der Norm gegen Null konvergieren, woraus $|\lambda_n| = \|A\varphi_n\| \to 0$ folgt.

Wir nehmen jetzt an, daß

$$M^{\perp} = H \ominus M(\varphi_1, \varphi_2, \ldots, \varphi_n, \ldots) = \bigcap_n M_n{}^{\perp} \neq 0$$

ist. Wenn dann ξ, $\xi \neq 0$, zu $M^{\perp}$ gehört, gilt $(A\xi, \xi) \leqq \lambda_n \|\xi\|^2$ für alle n, d. h.

$$(A\xi, \xi) = 0.$$

Wenden wir hier Lemma 3 (für $\max |(A\xi, \xi)| = 0$) auf $M^{\perp}$ an, so erhalten wir $A\xi = 0$. Also bildet der Operator A den Teilraum $M^{\perp}$ auf die Null ab.

Auf Grund der Konstruktion des Systems $\{\varphi_n\}$ ist klar, daß man jeden Vektor eindeutig in der Form

$$\xi = \sum c_k \varphi_k + \xi' \quad \text{mit} \quad A\xi' = 0$$

darstellen kann. Daraus ergibt sich

$$A\xi = \sum \lambda_k c_k \varphi_k.$$

Damit ist der Satz bewiesen.

Dieser Satz spielt eine fundamentale Rolle in der Theorie der Integralgleichungen, die in Kapitel 9 behandelt werden.

Bemerkung. Der soeben bewiesene Satz besagt, daß für jeden kompakten selbstadjungierten Operator A in H eine orthogonale Basis in H existiert, die aus Eigenwerten von A besteht. Um eine solche Basis zu erhalten, braucht man nur zu dem im Beweis konstruierten System von Eigenvektoren $\{\varphi_n\}$ noch eine beliebige orthogonale Basis des Teilraumes $M^\perp$ hinzuzufügen, die aus Eigenvektoren zum Eigenwert $\lambda = 0$ besteht. Mit anderen Worten: Wir haben für kompakte selbstadjungierte Operatoren im Hilbertraum ein vollkommen analoges Resultat erhalten wie für selbstadjungierte Operatoren in einem endlichdimensionalen Raum.

Für nichtselbstadjungierte Operatoren im n-dimensionalen Raum ist dieser Sachverhalt allgemein nicht richtig. Es gilt jedoch folgender Satz: *Jede lineare Abbildung im n-dimensionalen Raum besitzt wenigstens einen Eigenwert* (*und damit einen Eigenvektor*). Es ist nicht schwer zu beweisen, daß sich diese Behauptung nicht auf kompakte Operatoren in einem Hilbertraum überträgt. So ist zum Beispiel in l_2 der Operator A,

$$Ax = A(x_1, x_2, \ldots, x_n, \ldots) = \left(0, x_1, \frac{x_2}{2}, \ldots, \frac{x_{n-1}}{n-1}, \ldots\right), \tag{8}$$

kompakt, aber er besitzt keinen einzigen Eigenwert. (Den Beweis überlassen wir dem Leser.)

Aufgabe. Man bestimmte das Spektrum des Operators (8).

5. Maße, meßbare Funktionen, Integrale

Der Begriff des Maßes $\mu(A)$ einer Menge A ist eine natürliche Verallgemeinerung des Begriffes

1. der Länge $L(\Delta)$ einer Strecke Δ,

2. des Flächeninhaltes $S(F)$ eines ebenen Gebildes F,

3. des Volumens $V(G)$ eines räumlichen Gebildes G,

4. des Zuwachses $\varphi(b) - \varphi(a)$ einer nichtfallenden Funktion $\varphi(t)$ auf dem halboffenen Intervall $[a, b)$,

5. des Integrals einer nichtnegativen Funktion, erstreckt über irgendein lineares, ebenes oder räumliches Gebiet, usw.

Dieser Begriff entstand in der Theorie der Funktionen einer reellen Veränderlichen und ging von dort in die Wahrscheinlichkeitstheorie, die Theorie der dynamischen Systeme, die Funktionalanalysis und in viele andere Gebiete der Mathematik ein.

In 5.1. legen wir, ausgehend vom Begriff des Flächeninhalts eines Rechtecks, die Maßtheorie für Mengen in der Ebene dar. Die allgemeine Maßtheorie wird in 5.2. und 5.3. behandelt. Der Leser stellt leicht fest, daß alle in 5.1. durchgeführten Überlegungen allgemeinen Charakter haben und sich ohne wesentliche Veränderungen in der abstrakten Theorie wiederholen.

5.1. Das Maß ebener Mengen

5.1.1. Das Maß von Elementarmengen. Wir betrachten ein Mengensystem $\mathfrak{S}$ in der x, y-Ebene, dessen Mengen durch eine der Ungleichungen

$$a \leqq x \leqq b,$$
$$a < x \leqq b,$$
$$a \leqq x < b,$$
$$a < x < b$$

und eine der Ungleichungen

$$c \leqq y \leqq d,$$
$$c < y \leqq d,$$

$$c \leqq y < d,$$
$$c < y < d$$

bestimmt sind; a, b, c, d sind beliebige reelle Zahlen. Die Mengen, die zu diesem System gehören, nennen wir *Rechtecke*. Das abgeschlossene Rechteck, das durch die Ungleichungen

$$a \leqq x \leqq b, \qquad c \leqq y \leqq d$$

bestimmt wird, ist ein Rechteck im üblichen Sinn (mit Rand), wenn $a < b$ und $c < d$ ist, eine Strecke (wenn $a = b$ und $c < d$ oder wenn $a < b$ und $c = d$), ein Punkt (für $a = b$, $c = d$) und schließlich die leere Menge (wenn $a > b$ oder $c > d$). Das offene Rechteck

$$a < x < b, \qquad c < y < d$$

wird in Abhängigkeit von den Beziehungen zwischen a, b, c und d ein eigentliches Rechteck ohne Rand oder die leere Menge sein. Jedes Rechteck der restlichen Typen (wir nennen sie halboffen) ist ein eigentliches Rechteck ohne eine, zwei oder drei Randseiten, ein Intervall, ein Halbintervall oder die leere Menge.

Die Klasse aller Rechtecke in der x, y-Ebene bezeichnen wir mit $\mathfrak{S}$.

Für jedes dieser Rechtecke definieren wir das Maß mit Hilfe des aus der Elementargeometrie bekannten Flächeninhaltes von Rechtecken. Wir setzen fest:

a) Das Maß der leeren Menge ist gleich 0.

b) Das Maß eines nichtleeren, durch die Zahlen a, b, c und d definierten Rechtecks ist gleich

$$(b - a)\,(d - c).$$

Auf diese Weise wird jedem Rechteck P aus $\mathfrak{S}$ eine Zahl $m(P)$ — sein Maß — zugeordnet. Diese Zuordnung hat folgende Eigenschaften:

1. Das Maß $m(P)$ nimmt nur nichtnegative reelle Werte an.

2. Das Maß $m(P)$ ist additiv, d. h., wenn $P = \bigcup\limits_{k=1}^{n} P_k$ und $P_i \cap P_k = \emptyset$ für $i \neq k$ ist, dann gilt

$$m(P) = \sum_{k=1}^{n} m(P_k).$$

Unsere Aufgabe besteht nun darin, das bisher für Rechtecke definierte Maß $m(P)$ unter Erhaltung der Eigenschaften 1 und 2 auf eine umfangreichere Klasse von Mengen fortzusetzen.

Zuerst werden wir das Maß auf die sogenannten Elementarmengen erweitern. Wir nennen eine ebene Menge *Elementarmenge*, wenn sie wenigstens auf eine Weise als Vereinigung endlich vieler paarweise disjunkter Rechtecke dargestellt werden kann.

Im weiteren werden wir folgenden Satz benötigen:

Satz 1. *Vereinigung, Durchschnitt, Differenz und symmetrische Differenz zweier Elementarmengen sind wieder Elementarmengen.*

Das bedeutet, daß die Elementarmengen gemäß der in 1.5. eingeführten Terminologie einen *Ring* bilden.

Beweis. Es ist klar, daß der Durchschnitt zweier Rechtecke wieder ein Rechteck ist. Deshalb ist für die zwei Elementarmengen

$$A = \bigcup_k P_k \quad \text{und} \quad B = \bigcup_j Q_j$$

auch ihr Durchschnitt

$$A \cap B = \bigcup_{k,j} (P_k \cap Q_j)$$

eine Elementarmenge.

Die Differenz zweier Rechtecke ist, wie man leicht nachprüft, eine Elementarmenge. Bilden wir die Differenz zwischen einem Rechteck und einer Elementarmenge, so erhalten wir wieder eine Elementarmenge (als Durchschnitt von Elementarmengen). Es seien jetzt A und B Elementarmengen. Dann läßt sich offensichtlich ein Rechteck P finden, das jede dieser Mengen enthält. Die Menge

$$A \cup B = P \setminus [(P \setminus A) \cap (P \setminus B)]$$

ist dann nach den vorangehenden Betrachtungen eine Elementarmenge. Hieraus und aus den Gleichungen

$$A \setminus B = A \cap (P \setminus B),$$

$$A \triangle B = (A \cup B) \setminus (A \cap B)$$

folgt, daß die Differenz und die symmetrische Differenz von Elementarmengen wieder Elementarmengen sind. Damit ist der Satz bewiesen.

Wir definieren nun das Maß $m'(A)$ für Elementarmengen auf folgende Weise: Für

$$A = \bigcup_k P_k,$$

wobei die Mengen P_k paarweise disjunkte Rechtecke sind, sei

$$m'(A) = \sum_k m_k(P_k).$$

Wir zeigen, daß $m'(A)$ nicht von der Art der Zerlegung von A in endlich viele paarweise disjunkte Rechtecke abhängt. Es sei

$$A = \bigcup_k P_k = \bigcup_j Q_j,$$

wobei P_k und Q_j Rechtecke sind, für die $P_i \cap P_k = \emptyset$ und $Q_i \cap Q_k = \emptyset$ gilt, wenn $i \neq k$ ist. Weil der Durchschnitt $P_k \cap Q_j$ zweier Rechtecke ein Rechteck ist, folgt aus der Additivität des Maßes für Rechtecke

$$\sum_k m(P_k) = \sum_{k,j} m(P_k \cap Q_j) = \sum_j m(Q_j).$$

Insbesondere stimmt also für Rechtecke das Maß m' mit dem Ausgangsmaß m überein.

Es ist leicht zu sehen, daß das so definierte Maß für Elementarmengen nichtnegativ und additiv ist.

Wir stellen nun folgende wichtige Eigenschaft des Maßes von Elementarmengen fest.

Satz 2. *Wenn A eine Elementarmenge und $\{A_n\}$ ein endliches oder abzählbares System von Elementarmengen mit der Eigenschaft*

$$A \subset \bigcup_n A_n$$

ist, dann gilt

$$m'(A) \leqq \sum_n m'(A_n). \tag{1}$$

Beweis. Für ein beliebiges $\varepsilon > 0$ und eine vorgegebene Menge A kann man offensichtlich eine solche abgeschlossene Elementarmenge $\hat{A}$ finden, die in A enthalten ist und die Bedingung

$$m'(\hat{A}) \geqq m'(A) - \frac{\varepsilon}{2}$$

erfüllt. (Es genügt, jedes der k Rechtecke P_i aus einer Darstellung von A durch ein in ihm enthaltenes abgeschlossenes Rechteck mit einer Fläche größer als $m(P_i) - \frac{\varepsilon}{2k}$ zu ersetzen.)

Weiter kann man für jede Menge A_n eine offene Elementarmenge $\tilde{A}_n$ finden, die A_n enthält und der Bedingung

$$m'(\tilde{A}_n) \leqq m'(A_n) + \frac{\varepsilon}{2^{n+1}}$$

genügt, Es ist klar, daß

$$\hat{A} \subset \bigcup_n \tilde{A}_n$$

gilt. Aus dem System $\{\tilde{A}_n\}$ kann man (nach dem Satz von Heine-Borel) ein endliches System $\tilde{A}_{n_1}, \ldots, \tilde{A}_{n_s}$ auswählen, das $\hat{A}$ ebenfalls überdeckt. Dabei gilt offensichtlich

$$m'(\hat{A}) \leqq \sum_{i=1}^{s} m'(\tilde{A}_{n_i})$$

(weil sonst $\hat{A}$ durch endlich viele offene Rechtecke überdeckt werden würde, deren Gesamtfläche kleiner als $m'(\hat{A})$ wäre, was natürlich nicht möglich ist). Daher ist

$$m'(A) \leqq m'(\hat{A}) + \frac{\varepsilon}{2} \leqq \sum_{i=1}^{s} m'(\tilde{A}_{n_i}) + \frac{\varepsilon}{2} \leqq \sum_{n} m'(\tilde{A}_n) + \frac{\varepsilon}{2}$$

$$\leqq \sum_{n} m'(A_n) + \sum_{n} \frac{\varepsilon}{2^{n+1}} + \frac{\varepsilon}{2} = \sum_{n} m'(A_n) + \varepsilon,$$

woraus (1) folgt, da $\varepsilon > 0$ beliebig gewählt werden konnte. Damit ist der Satz bewiesen.

Die im Satz 2 festgestellte Eigenschaft des Maßes m' (das Maß einer Menge ist nicht größer als die Summe der Maße von endlich oder abzählbar unendlich vielen offenen Mengen, die die betrachtete Menge überdecken) bezeichnen wir als *Subadditivität.* Aus ihr folgt die Eigenschaft der sogenannten *abzählbaren Additivität* oder *σ-Additivität,* die in folgendem besteht:

Es sei eine Elementarmenge A als Vereinigung abzählbar vieler paarweise disjunkter Elementarmengen A_n $(n = 1, 2, \ldots)$,

$$A = \bigcup_{n=1}^{\infty} A_n,$$

dargestellt. Dann ist

$$m'(A) = \sum_{n=1}^{\infty} m'(A_n)$$

(d. h., das Maß der Vereinigung abzählbar vieler paarweise disjunkter Elementarmengen ist gleich der Summe der Maße der Elementarmengen).

Auf Grund der Additivität des Maßes m' gilt nämlich für jedes N

$$m'(A) \geqq m'\left(\bigcup_{n=1}^{N} A_n\right) = \sum_{n=1}^{N} m'(A_n),$$

und wenn wir für $N \to \infty$ zum Grenzwert übergehen, erhalten wir

$$m'(A) \geqq \sum_{n=1}^{\infty} m'(A_n).$$

Nach Satz 2 gilt auch die entgegengesetzte Ungleichung. Damit ist die σ-Additivität des Maßes m' bewiesen.

Bemerkung. Beim Leser kann der Eindruck entstehen, daß man die σ-Additivität des Maßes in der Ebene automatisch aus seiner Additivität durch Grenzübergang erhält. In Wirklichkeit ist das jedoch nicht so (im Beweis von Satz 2 benutzen wir den Satz von HEINE-BOREL, der auf der Verknüpfung von metrischen und topo-

logischen Eigenschaften ebener Mengen beruht). In 5.2. werden wir bei der Untersuchung von Maßen auf beliebigen abstrakten Mengen sehen, daß aus der Additivität des Maßes im allgemeinen nicht seine σ-Additivität folgt.

5.1.2. Das Lebesguesche Maß ebener Mengen. Die Elementarmengen schöpfen die Klasse der für die Geometrie und die klassische Analysis interessanten Mengen nicht aus. Daher entsteht die Frage, ob man den Begriff des Maßes unter Erhaltung seiner Grundeigenschaften auf eine umfangreichere Klasse als die der endlichen Vereinigungen von achsenparallelen Rechtecken ausdehnen kann.

Eine im angegebenen Sinn abschließende Lösung dieses Problems wurde von H. Lebesgue am Anfang des 20. Jahrhunderts gegeben.

Bei der Darlegung der Lebesgueschen Maßtheorie werden wir nicht nur endliche, sondern auch unendliche Vereinigungen von Rechtecken betrachten müssen. Um nicht gleich mit Mengen „unendlichen Maßes" in Konflikt zu kommen, beschränken wir uns deshalb zuerst auf Mengen, die ganz im Quadrat $E = \{0 \leqq x \leqq 1; 0 \leqq y \leqq 1\}$ liegen.

Auf der Gesamtheit aller dieser Mengen definieren wir auf folgende Weise eine Funktion $\mu^*(A)$.

Definition 1. Jeder Menge A wird die Zahl

$$\mu^*(A) = \inf_{A \subset \bigcup_k P_k} \sum_k m(P_k), \tag{1}$$

als *äußeres Maß* zugeordnet. Dabei wird das Infimum über alle möglichen Überdeckungen der Menge A durch endliche oder abzählbare Systeme von Rechtecken gebildet.

Bemerkungen

1. Wenn wir in der Definition des äußeren Maßes statt der Überdeckungen durch Recktecke Überdeckungen durch beliebige Elementarmengen (endlich oder abzählbar viele) nehmen, so erhalten wir offensichtlich denselben Wert für $\mu^*(A)$, weil jede Elementarmenge ihrerseits eine endliche Vereinigung von Rechtecken ist.

2. Wenn A Elementarmenge ist, dann ist $\mu^*(A) = m'(A)$. Denn sind $P_1, \ldots, P_n$ die Rechtecke einer Darstellung von A, dann ist nach Definition

$$m'(A) = \sum_{i=1}^{n} m(P_i).$$

Weil die Rechtecke P_i die Menge A überdecken, ist

$$\mu^*(A) \leqq \sum_{i=1}^{n} m(P_i) = m'(A).$$

Andererseits gilt nach Satz 2 für ein beliebiges (endliches oder abzählbares) System $\{Q_j\}$ von Rechtecken, das A überdeckt, $m'(A) \leqq \sum_j m(Q_j)$, insgesamt gilt also $\mu^*(A) = m'(A)$.

Satz 3. *Ist $A \subset \bigcup_n A_n$, wobei $\{A_n\}$ ein endliches oder abzählbares Mengensystem ist, dann gilt*

$$\mu^*(A) \leqq \sum_n \mu^*(A_n). \tag{2}$$

Insbesondere ist $\mu^*(A) \leqq \mu^*(B)$ für $A \subset B$.

Beweis. Nach Definition des äußeren Maßes läßt sich für jedes A_n ein solches System $\{P_{nk}\}$ von Rechtecken (endlich oder abzählbar) finden, daß $A_n \subset \bigcup_k P_{nk}$ und

$$\sum_k m(P_{nk}) \leqq \mu^*(A_n) + \frac{\varepsilon}{2^n}$$

gilt, wobei $\varepsilon > 0$ beliebig vorgegeben werden kann. Daraus folgt

$$A \subset \bigcup_n \bigcup_k P_{nk}$$

und

$$\mu^*(A) \leqq \sum_n \sum_k m(P_{nk}) \leqq \sum_n \mu^*(A_n) + \varepsilon$$

für jedes $\varepsilon > 0$. Das ist gleichbedeutend mit der Aussage des Satzes.

Da m' und μ^* auf den Elementarmengen übereinstimmen, wird Satz 2 nun zu einem Spezialfall von Satz 3.

Definition 2. Die Menge A heißt *meßbar* (im Lebesgueschen Sinne), wenn sich für jede Zahl $\varepsilon > 0$ eine solche Elementarmenge B finden läßt, daß

$$\mu^*(A \triangle B) < \varepsilon \tag{3}$$

ist. Die Einschränkung der Funktion μ^* auf die meßbaren Mengen heißt *Lebesguesches Maß* und wird im weiteren mit μ bezeichnet.

Bemerkung. Die eben eingeführte Definition besitzt einen anschaulichen Sinn. Sie besagt nämlich, daß eine Menge genau dann meßbar ist, wenn man sie beliebig gut bezüglich des äußeren Maßes durch Elementarmengen approximieren kann.

Nachdem wir die Klasse $\mathfrak{M}_E$ der (im Lebesgueschen Sinne) meßbaren Mengen und eine Funktion μ, das Lebesguesche Maß, auf dieser Klasse definiert haben, besteht unser nächstes Ziel darin, folgende Aussagen zu begründen.

1. *Die Gesamtheit $\mathfrak{M}_E$ der meßbaren Mengen ist abgeschlossen bezüglich der Bildung von endlich oder abzählbar vielen Vereinigungen und Durchschnitten* (d. h., sie ist eine σ-Algebra; vgl. die Definition in 1.5.4.)

2. *Die Funktion μ ist σ-additiv auf $\mathfrak{M}_E$.*

Die folgenden Sätze sind Etappen des Beweises dieser Behauptungen.

Satz 4. *Das Komplement einer meßbaren Menge ist meßbar.*

Das folgt sofort aus der Identität

$$(E \setminus A) \triangle (E \setminus B) = A \triangle B,$$

die ihrerseits unmittelbar klar ist.

Satz 5. *Vereinigung und Durchschnitt endlich vieler meßbarer Mengen sind wieder meßbare Mengen.*

Beweis. Es genügt offensichtlich, die Aussage für zwei Mengen zu beweisen. A_1 und A_2 seien meßbare Mengen. Dann lassen sich für jedes $\varepsilon > 0$ solche Elementarmengen B_1 und B_2 finden, daß

$$\mu^*(A_1 \triangle B_1) < \frac{\varepsilon}{2}, \qquad \mu^*(A_2 \triangle B_2) < \frac{\varepsilon}{2}$$

ist. Weil

$$(A_1 \cup A_2) \triangle (B_1 \cup B_2) \subset (A_1 \triangle B_1) \cup (A_2 \triangle B_2)$$

ist, folgt

$$\mu^*[(A_1 \cup A_2) \triangle (B_1 \cup B_2)] \leqq \mu^*(A_1 \triangle B_1) + \mu^*(A_2 \triangle B_2) < \varepsilon.$$

Da $B_1 \cup B_2$ eine Elementarmenge ist, bedeutet diese Aussage die Meßbarkeit von $A_1 \cup A_2$.

Die Meßbarkeit des Durchschnitts zweier meßbarer Mengen folgt aus Satz 4 und dem bisherigen Beweis, wenn man von der Beziehung

$$A_1 \cap A_2 = E \setminus [(E \setminus A_1) \cup (E \setminus A_2)] \tag{4}$$

ausgeht.

Folgerung. *Die Differenz und die symmetrische Differenz zweier meßbarer Mengen sind meßbar.*

Das folgt aus den Sätze 4 und 5 und den Identitäten

$$A_1 \setminus A_2 = A_1 \cap (E \setminus A_2), \qquad A_1 \triangle A_2 = (A_1 \setminus A_2) \cup (A_2 \setminus A_1).$$

Satz 6. *Wenn $A_1, \ldots, A_n$ paarweise disjunkte meßbare Mengen sind, dann gilt*

$$\mu\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mu(A_k). \tag{5}$$

Für den Beweis dieses Satzes brauchen wir folgendes Lemma.

Lemma. *Für zwei beliebige Mengen A und B ist*

$$|\mu^*(A) - \mu^*(B)| \leqq \mu^*(A \triangle B).$$

Beweis des Lemmas. Wegen

$$A \subset B \cup (A \triangle B)$$

folgt aus Satz 3

$$\mu^*(A) \leqq \mu^*(B) + \mu^*(A \triangle B).$$

Hieraus folgt die Behauptung im Fall $\mu^*(A) \geqq \mu^*(B)$. Ist aber $\mu^*(A) \leqq \mu^*(B)$, dann folgt die Behauptung des Lemmas aus der Ungleichung

$$\mu^*(B) \leqq \mu^*(A) + \mu^*(A \triangle B),$$

die analog gewonnen werden kann.

Beweis von Satz 6. Wie im Satz 5 genügt es, zwei Mengen zu betrachten. Wir wählen ein beliebiges $\varepsilon > 0$ und zwei solche Elementarmengen B_1 und B_2, daß

$$\mu^*(A_1 \triangle B_1) < \varepsilon, \tag{6}$$

$$\mu^*(A_2 \triangle B_2) < \varepsilon. \tag{7}$$

Setzen wir $A = A_1 \cup A_2$ und $B = B_1 \cup B_2$, so ist die Menge A meßbar nach Satz 5. Weil sich die Mengen A_1 und A_2 nicht schneiden, ist

$$B_1 \cap B_2 \subset (A_1 \triangle B_1) \cup (A_2 \triangle B_2)$$

und folglich

$$m'(B_1 \cap B_2) \leqq 2\varepsilon. \tag{8}$$

Nach dem Lemma folgt aus (6) und (7)

$$|m'(B_1) - \mu^*(A_1)| < \varepsilon, \tag{9}$$

$$|m'(B_2) - \mu^*(A_2)| < \varepsilon. \tag{10}$$

Da auf der Gesamtheit der Elementarmengen das Maß additiv ist, erhalten wir aus (8) bis (10)

$$m'(B) = m'(B_1) + m'(B_2) - m'(B_1 \cap B_2) \geqq \mu^*(A_1) + \mu^*(A_2) - 4\varepsilon.$$

Gehen wir nun von der Inklusion $B \subset A \cup (A \triangle B)$ aus und beachten die Beziehung $A \triangle B \subset (A_1 \triangle B_1) \cup (A_2 \triangle B_2)$ sowie (6), (7) und die vorhergehende Ungleichung, so können wir abschätzen:

$$\mu^*(A) \geqq m'(B) - \mu^*(A \triangle B) \geqq m'(B) - 2\varepsilon \geqq \mu^*(A_1) + \mu^*(A_2) - 6\varepsilon.$$

Da wir $\varepsilon > 0$ beliebig klein wählen können, ist

$$\mu^*(A) \geqq \mu^*(A_1) + \mu^*(A_2).$$

Nach Satz 3 ist die umgekehrte Ungleichung

$$\mu^*(A) \leqq \mu^*(A_1) + \mu^*(A_2)$$

immer richtig. Insgesamt gilt also

$$\mu^*(A) = \mu^*(A_1) + \mu^*(A_2).$$

Weil A_1, A_2 und A meßbar sind, kann man in dieser Gleichung μ^* durch μ ersetzen. Damit ist der Satz bewiesen.

Aus diesem Satz folgt insbesondere für jede meßbare Menge A

$$\mu(E \setminus A) = 1 - \mu(A).$$

Satz 7. *Vereinigung und Durchschnitt abzählbar vieler meßbarer Mengen sind wieder meßbare Mengen.*

Beweis. Es sei $A_1, A_2, \ldots, A_n, \ldots$ ein abzählbares System meßbarer Mengen und $A = \bigcup_{n=1}^{\infty} A_n$. Setzen wir $A_n' = A_n \setminus \bigcup_{k=1}^{n-1} A_k$, dann ist wieder $A = \bigcup_{n=1}^{\infty} A_n'$, und die Mengen A_n' sind paarweise disjunkt. Nach Satz 5 und der Folgerung aus ihm sind alle Mengen A_n' meßbar. Nach Satz 6 und der Definition des äußeren Maßes ist für jede endliche Zahl n

$$\sum_{k=1}^{n} \mu(A_k') = \mu\left(\bigcup_{k=1}^{n} A_k'\right) \leqq \mu^*(A).$$

Daraus folgt, daß die Reihe

$$\sum_{n=1}^{\infty} \mu(A_n')$$

konvergiert. Also läßt sich zu jedem $\varepsilon > 0$ eine solche Zahl N finden, daß

$$\sum_{n>N} \mu(A_n') < \frac{\varepsilon}{2} \tag{11}$$

ist. Die Menge $C = \bigcup_{n=1}^{N} A_n'$ ist meßbar (als Vereinigung endlich vieler meßbarer Mengen). Damit gibt es für sie eine Elementarmenge B mit

$$\mu^*(C \triangle B) < \frac{\varepsilon}{2}. \tag{12}$$

Aus

$$A \triangle B \subset (C \triangle B) \cup \Big(\bigcup_{n>N} A_n'\Big)$$

und aus (11) und (12) folgt

$$\mu^*(A \triangle B) < \varepsilon,$$

d. h., A ist meßbar.

Weil das Komplement meßbarer Mengen meßbar ist, folgt die Behauptung des Satzes bezüglich des Durchschnitts aus der Identität

$$\bigcap_n A_n = E \setminus \bigcup_n (E \setminus A_n).$$

Der Satz 7 ist stärker als der Satz 5. Der folgende Satz stellt eine entsprechende Verstärkung des Satzes 6 dar.

Satz 8. *Wenn $\{A_n\}$ eine Folge paarweise disjunkter meßbarer Mengen und $A = \bigcup_n A_n$ ist, gilt*

$$\mu(A) = \sum_n \mu(A_n).$$

Beweis. Nach Satz 6 ist für jedes N

$$\mu\Big(\bigcup_{n=1}^{N} A_n\Big) = \sum_{n=1}^{N} \mu(A_n) < \mu(A).$$

Gehen wir für $N \to \infty$ zum Grenzwert über, so erhalten wir

$$\sum_{n=1}^{\infty} \mu(A_n) \leqq \mu(A). \tag{13}$$

Andererseits ist nach Satz 3

$$\mu(A) \leqq \sum_{n=1}^{\infty} \mu(A_n). \tag{14}$$

Aus (13) und (14) folgt die Behauptung des Satzes.

Die in Satz 8 formulierte Eigenschaft des Maßes wird als *abzählbare Additivität* oder *σ-Additivität* des Maßes bezeichnet. Aus der σ-Additivität ergibt sich die folgende Eigenschaft des Maßes, die als *Stetigkeit* des Maßes bezeichnet wird.

Satz 9. *Wenn $A_1 \supset A_2 \supset \cdots$ eine Folge ineinanderliegender meßbarer Mengen und $A = \bigcap_n A_n$ ist, gilt*

$$\mu(A) = \lim_{n\to\infty} \mu(A_n).$$

Beweis. Es genügt, den Fall $A = \emptyset$ zu betrachten, da der allgemeine Fall auf diesen mit der Substitution von A_n durch $A_n \setminus A$ zurückgeführt werden kann. Es

ist

$$A_1 = (A_1 \setminus A_2) \cup (A_2 \setminus A_3) \cup \cdots$$

und

$$A_n = (A_n \setminus A_{n+1}) \cup (A_{n+1} \setminus A_{n+2}) \cup \ldots,$$

wobei die Mengen $A_i \setminus A_{i+1}$ für verschiedene i disjunkt sind. Daher folgt aus der σ-Additivität des Maßes μ

$$\mu(A_1) = \sum_{k=1}^{\infty} \mu(A_k \setminus A_{k+1}) \tag{15}$$

und

$$\mu(A_n) = \sum_{k=n}^{\infty} \mu(A_k \setminus A_{k+1}). \tag{16}$$

Weil die Reihe (15) konvergiert, muß ihr Rest (16) für $n \to \infty$ gegen 0 streben, d. h., es ist

$$\mu(A_n) \to 0 \quad \text{für} \quad n \to \infty,$$

was zu zeigen war.

Folgerung. *Wenn* $A_1 \subset A \subset \cdots$ *eine wachsende Folge meßbarer Mengen und* $A = \bigcup_n A_n$ *ist, gilt*

$$\mu(A) = \lim_{n\to\infty} \mu(A_n).$$

Zum Beweis geht man von den Mengen A_n zu ihren Komplementen $E \setminus A_n$ über und benutzt Satz 9.

Zum Abschluß weisen wir noch auf eine offensichtliche, aber wichtige Tatsache hin. *Jede Menge A, deren äußeres Maß gleich* 0 *ist, ist meßbar*. Zur Begründung genügt es, in der Definition der Meßbarkeit $B = \emptyset$ zu setzen, dann ist nämlich

$$\mu^*(A \triangle B) = \mu^*(A \triangle \emptyset) = \mu^*(A) = 0 < \varepsilon.$$

Fassen wir zusammen. In diesem Abschnitt wurde das Maß von den Elementarmengen auf die umfangreichere Klasse $\mathfrak{M}_E$ ausgedehnt, die bezüglich der Bildung von abzählbar vielen Vereinigungen und Durchschnitten abgeschlossen, d. h. eine σ-Algebra ist. Das konstruierte Maß ist auf dieser Klasse σ-additiv. Die oben angegebenen Sätze erlauben folgende Aussagen über die Gesamtheit der Lebesgue-meßbaren Mengen.

Jede offene Menge, die in E gelegen ist, kann als Vereinigung endlich oder abzählbar vieler offener Rechtecke, d. h. meßbarer Mengen, dargestellt werden. Nach Satz 7 sind damit alle offenen Mengen meßbar. Die abgeschlossenen Mengen sind die Komplemente offener Mengen und folglich auch meßbar. Aus Satz 7 folgt nun die Meßbarkeit aller Mengen, die man aus den offenen und abgeschlossenen Mengen mit Hilfe endlich oder abzählbar vieler Vereinigungen oder Durchschnitte erhalten

kann. Später wird gezeigt, daß sich die Gesamtheit der meßbaren Mengen in diesen Mengen nicht erschöpft.

5.1.3. Einige Ergänzungen und Verallgemeinerungen. Bis jetzt haben wir nur Mengen betrachtet, welche im Einheitsquadrat

$$E = \{0 \leqq x \leqq 1, 0 \leqq y \leqq 1\}$$

enthalten waren. Wir können uns leicht von dieser Einschränkung befreien, zum Beispiel auf folgende Weise. Man stellt die ganze Ebene als Vereinigung halboffener Quadrate

$$E_{nm} = \{n < x \leqq n + 1, m < y \leqq m + 1\}$$

(n, m ganze Zahlen) dar und setzt fest, daß die ebene Menge A meßbar ist, wenn ihr Durchschnitt $A_{nm} = A \cap E_{nm}$ mit jedem dieser Quadrate meßbar im obigen Sinne ist. Dabei setzen wir

$$\mu(A) = \sum_{n,m} \mu(A_{nm})$$

durch Definition fest. Die rechtsstehende Reihe konvergiert entweder gegen einen endlichen Wert oder divergiert bestimmt gegen $+\infty$. Daher kann das Maß μ jetzt auch unendliche Werte annehmen. Alle Eigenschaften des Maßes und der meßbaren Mengen, die oben angegeben wurden, lassen sich offensichtlich auf den jetzt betrachteten allgemeineren Fall übertragen. Man muß nur darauf achten, daß die Vereinigung abzählbar vieler meßbarer Mengen mit endlichem Maß selbst ein unendliches Maß haben kann. Die Klasse der meßbaren Mengen der ganzen Ebene bezeichnen wir mit $\mathfrak{A}$.

In 5.1.1. und 5.1.2. haben wir die Konstruktion des Lebesgueschen Maßes für ebene Mengen dargestellt. Analog kann man das Lebesguesche Maß auf der Geraden, im dreidimensionalen Raum oder allgemein im euklidischen Raum beliebiger Dimension n konstruieren. In jedem dieser Fälle konstruiert man das Maß auf ein und dieselbe Weise: Ausgehend von einem Maß, das auf irgendeinem System einfacher Mengen (Rechtecke im Fall der Ebene; Intervalle (a, b), Segmente $[a, b]$ und Halbsegmente $(a, b]$, $[a, b)$ im Fall der Geraden; usw.) gegeben ist, definieren wir das Maß zuerst für endliche Vereinigungen solcher Mengen und dehnen es danach auf eine wesentlich breitere Klasse von Mengen, die Lebesgue-meßbaren Mengen aus. Die Definition der Meßbarkeit selbst überträgt sich wörtlich auf Mengen in einem Raum beliebiger Dimension.

Bisher haben wir den Begriff des Lebesgueschen Maßes ausgehend vom üblichen Flächeninhalt konstruiert. Die analoge Konstruktion für den eindimensionalen Fall stützt sich auf den Begriff der Länge eines Intervalls (Segments, Halbsegments). Man kann den Begriff des Maßes auch anders, auf eine allgemeinere Weise einführen.

Es sei $F(t)$ irgend eine nichtfallende, von links stetige Funktion auf der Geraden. Wir definieren

$$m\big((a, b)\big) = F(b) - F(a + 0),$$
$$m\big([a, b]\big) = F(b + 0) - F(a),$$
$$m\big((a, b]\big) = F(b + 0) - F(a + 0),$$
$$m\big([a, b)\big) = F(b) - F(a).$$

Es ist leicht zu sehen, daß die so definierte Intervallfunktion m nichtnegativ und additiv ist. Führen wir für diese Funktion analoge Überlegungen wie im vorhergehenden durch, so kommen wir zu einem Maß μ_F. Dabei ist die Gesamtheit $\mathfrak{A}_F$ der bezüglich dieses Maßes meßbaren Mengen wieder bezüglich der Bildung von abzählbar vielen Vereinigungen und Durchschnitten abgeschlossen. Das Maß μ_F ist σ-additiv. Die Klasse $\mathfrak{A}_F$ der Mengen, die bezüglich μ_F meßbar sind, wird im allgemeinen von der Wahl der Funktion F abhängen. Die offenen und abgeschlossenen Mengen und folglich auch ihre abzählbaren Vereinigungen und Durchschnitte sind jedoch für jede Funktion F (mit den angegebenen Eigenschaften) meßbar.

Maße, die man auf diese Weise erhält, nennt man *Lebesgue-Stieltjessche Maße.* Insbesondere entspricht der Funktion $F(t) \equiv t$ das gewöhnliche Lebesguesche Maß der Geraden.

Wenn das Maß μ_F die Eigenschaft besitzt, daß es gleich 0 ist für jede Menge mit dem Lebesgueschen Maß 0, dann heißt das Maß μ_F *absolut stetig* (bezüglich μ). Wenn das Maß μ_F vollständig auf einer endlichen oder abzählbaren Punktmenge konzentriert ist (das ist dann der Fall, wenn die Menge der Funktionswerte von F endlich oder abzählbar ist), so heißt das Maß μ_F *diskret.* Das Maß μ_F heißt *singulär,* wenn es für jede einpunktige Menge gleich 0 ist und eine Menge mit dem Lebesgueschen Maß 0 existiert, für deren Komplement μ_F gleich 0 ist.

Man kann zeigen, daß jedes Maß μ_F als Summe eines absolutstetigen, eines diskreten und eines singulären Maßes darstellbar ist. Zu den Lebesgue-Stieltjesschen Maßen werden wir in Kapitel 6 zurückkehren.

Existenz nichtmeßbarer Mengen. Wir haben gesehen, daß die Klasse der Lebesguemeßbaren Mengen sehr umfangreich ist. Es entsteht die Frage, ob überhaupt nichtmeßbare Mengen existieren. Diese Frage muß positiv beantwortet werden. Am einfachsten sind nichtmeßbare Mengen auf der Kreislinie zu konstruieren, wenn dort das Lebesguesche Maß betrachtet wird.

Es sei C ein Kreis mit dem Umfang 1 und α eine irrationale Zahl. Die Menge der Punkte der Kreislinie teilen wir folgendermaßen in Klassen ein. Wir fassen alle diejenigen Punkte in einer Klasse zusammen, die durch eine Drehung des Kreises C um den Winkel $n\alpha\pi$ (n ganze Zahl) ineinander übergeführt werden können. Jede dieser Klassen besteht offensichtlich aus abzählbar vielen Punkten. Wählen wir aus jeder dieser Klassen einen Punkt aus und fassen diese Punkte zu einer Menge Φ_0 zusammen, so ist diese Menge Φ_0, wie im folgenden gezeigt wird, nicht meßbar.

Bezeichnen wir mit Φ_n die Menge, die aus Φ_0 durch eine Drehung um den Winkel $n\alpha\pi$ entsteht, so ist leicht zu sehen, daß sich die Mengen Φ_n paarweise nicht schneiden und die Vereinigung aller dieser Mengen gleich dem ganzen Kreis C ist. Wenn die Menge Φ_0 meßbar wäre, dann wären auch die zu ihr kongruenten Mengen Φ_n meßbar. Wegen

$$C = \bigcup_{n=-\infty}^{\infty} \Phi_n, \qquad \Phi_n \cap \Phi_m = \emptyset \quad \text{für} \quad n \neq m,$$

würde auf Grund der σ-Additivität des Maßes hieraus

$$1 = \sum_{n=-\infty}^{\infty} \mu(\Phi_n) \tag{17}$$

folgen. Da kongruente Mengen ein und dasselbe Maß haben müssen, ergibt sich andererseits aus der Meßbarkeit von Φ_0

$$\mu(\Phi_n) = \mu(\Phi_0).$$

Offensichtlich ist diese Relation ein Widerspruch zu (17). Denn die Summe der Reihe der rechten Seite von (17) ist gleich 0, wenn $\mu(\Phi_0) = 0$ ist, und gleich ∞, wenn $\mu(\Phi_0) > 0$ ist. Dieser Widerspruch zeigt, daß die Menge Φ_0 (und damit auch jede Menge Φ_n) nicht meßbar ist.

5.2. Der allgemeine Maßbegriff. Fortsetzung eines Maßes von einem Semiring auf einen Ring. Additivität und σ-Additivität[1]

5.2.1. Definition des Maßes. Bisher haben wir, ausgehend vom Maß (Flächeninhalt) von Rechtecken, durch Fortsetzung dieses Maßes auf eine umfangreichere Klasse von Mengen einen Maßbegriff für ebene Mengen konstruiert. Dabei war für unsere Konstruktion nicht der konkrete Ausdruck des Flächeninhalts von Rechtecken wesentlich, sondern allein seine allgemeinen Eigenschaften. So benutzten wir bei der Fortsetzung des ebenen Maßes von den Rechtecken auf die Elementarmengen nur, daß der Flächeninhalt eine nichtnegative additive Mengenfunktion und die Gesamtheit der Rechtecke ein Semiring ist. Bei der Konstruktion der Lebesgueschen Fortsetzung des ebenen Maßes war außerdem noch die σ-Additivität der Mengenfunktion wichtig. Daher kann man der in 5.1. für ebene Mengen durchgeführten Konstruktion eine abstrakte Form geben, die eine wesentlich breitere Anwendungsmöglichkeit besitzt. Diese abstrakte Konstruktion eines Maßes wird im weiteren beschrieben.

Zuerst geben wird folgende grundlegende Definition.

Definition 1. Eine Mengenfunktion $\mu(A)$ heißt *Maß*, wenn

1. das Definitionsgebiet $\mathfrak{S}_\mu$ der Funktion $\mu(A)$ ein Semiring ist,

[1] Hier und im folgenden werden wir systematisch die Begriffe und Fakten aus 1.5. benutzen.

2. die Werte der Funktion $\mu(A)$ reell und nichtnegativ sind,

3. $\mu(A)$ additiv ist, d. h. für eine beliebige endliche Zerlegung der Menge $A \in \mathfrak{S}_\mu$ in paarweise disjunkte Mengen $A_k \in \mathfrak{S}_\mu$,

$$A = \bigcup_{k=1}^{n} A_k,$$

die Beziehung

$$\mu(A) = \sum_{k=1}^{n} \mu(A_k)$$

gilt.

Bemerkung. Aus 3. und der Zerlegung $\emptyset = \emptyset \cup \emptyset$ folgt sofort $\mu(\emptyset) = 2\mu(\emptyset)$, also $\mu(\emptyset) = 0$.

5.2.2. Fortsetzung eines Maßes von einem Semiring auf den von ihm erzeugten Ring. Bei der Konstruktion des Maßes ebener Mengen wurde im ersten Schritt das Maß von den Rechtecken auf die Elementarmengen, d. h. auf endliche Vereinigungen paarweise disjunkter Rechtecke fortgesetzt. Jetzt geben wir das abstrakte Analogon dieser Konstruktion an. Zur Präzisierung des Begriffes der Fortsetzung definieren wir:

Definition 2. Das Maß μ heißt *Fortsetzung* des Maßes m, wenn $\mathfrak{S}_m \subset \mathfrak{S}_\mu$ und für jede Menge $A \in \mathfrak{S}_m$

$$\mu(A) = m(A)$$

ist.

Satz 1. *Für jedes auf einem Semiring $\mathfrak{S}_m$ gegebene Maß $m(A)$ existiert eine eindeutige bestimmte Fortsetzung $m'(A)$ auf den von $\mathfrak{S}_m$ erzeugten Ring $\mathfrak{R}(\mathfrak{S}_m)$, d. h. $\mathfrak{S}_{m'} = \mathfrak{R}(\mathfrak{S}_m)$.*

Beweis. Nach Satz 3 aus 1.5. existiert für jede Menge $A \in \mathfrak{R}(\mathfrak{S}_m)$ eine Zerlegung

$$A = \bigcup_{k=1}^{n} B_k \qquad (B_k \in \mathfrak{S}_m,\ B_k \cap B_l = \emptyset \quad \text{für} \quad k \neq l). \tag{1}$$

Wir setzen

$$m'(A) = \sum_{k=1}^{n} m(B_k) \tag{2}$$

und zeigen, daß diese Funktion die im Satz beschriebene Fortsetzung des Maßes m ist.

Zuerst weisen wir nach, daß die durch (2) definierte Größe $m'(A)$ unabhängig von der Wahl der Zerlegung (1) ist. Sind zwei Zerlegungen der Form (1) für A gegeben,

$$A = \bigcup_{i=1}^{n} B_i = \bigcup_{j=1}^{r} C_j, \qquad B_i \in \mathfrak{S}_m, \qquad C_j \in \mathfrak{S}_m,$$

so gehören alle Durchschnitte $B_i \cap C_j$ auch zum Semiring $\mathfrak{S}_m$ und sind paarweise disjunkt. Aus den offensichtlich geltenden Zerlegungen

$$B_i = \bigcup_{j=1}^{r} (B_i \cap C_j), \qquad C_j = \bigcup_{i=1}^{n} (B_i \cap C_j)$$

und der Additivität des Maßes m folgt nun

$$\sum_{i=1}^{n} m(B_i) = \sum_{i=1}^{n} \sum_{j=1}^{r} m(B_i \cap C_j) = \sum_{j=1}^{r} m(C_j),$$

was zu zeigen war.

Daß die durch (2) definierte Funktion nichtnegativ und additiv ist, ergibt sich unmittelbar aus den Eigenschaften von m. Damit ist die Existenz einer Fortsetzung m' des Maßes m auf den Ring $\mathfrak{R}(\mathfrak{S}_m)$ bewiesen.

Die Eindeutigkeit der Fortsetzung ergibt sich folgendermaßen. Ist $A \in \mathfrak{R}(\mathfrak{S}_m)$ und $A = \bigcup_{k=1}^{n} B_k$ eine Zerlegung der Form (1), dann gilt für jede Fortsetzung $\tilde{m}$ von m auf $\mathfrak{R}(\mathfrak{S}_m)$

$$\tilde{m}(A) = \sum_{k=1}^{n} \tilde{m}(B_k) = \sum_{k=1}^{n} m(B_k) = m'(A),$$

d. h., $\tilde{m}$ stimmt mit dem durch (2) definierten Maß m' überein. Damit ist der Satz vollständig bewiesen.

In diesem Beweis haben wir in abstrakten Termini eigentlich genau das Verfahren wiederholt, nach dem wir in 5.1. das Maß von den Rechtecken auf die Elementarmengen fortgesetzt hatten. Die Klasse der Elementarmengen ist dabei gerade der minimale Ring, der vom Semiring der Rechtecke erzeugt wird.

Aus der Additivität und der Nichtnegativität des Maßes ergeben sich folgende offensichtliche, aber doch wichtige Eigenschaften.

Satz 2. *Ist m ein auf einem beliebigen Ring $\mathfrak{R}_m$ gegebenes Maß und gehören die Mengen $A, A_1, \ldots, A_n$ zu $\mathfrak{R}_m$, dann folgt*

1. *aus $\bigcup_{k=1}^{n} A_k \subset A$ und $A_i \cap A_j = \emptyset$ für $i \neq j$ die Ungleichung*

$$\sum_{k=1}^{n} m(A_k) \leqq m(A);$$

2. *aus $\bigcup_{k=1}^{n} A_k \supset A$ die Ungleichung*

$$\sum_{k=1}^{n} m(A_k) \geqq m(A).$$

Insbesondere ist für $A \subset A'$ und $A, A' \in \mathfrak{R}_m$

$$m(A) \leqq m(A').$$

Beweis. Sind die Mengen $A_1, \ldots, A_n$ paarweise disjunkt und alle in A enthalten, so ist auf Grund der Additivität des Maßes

$$m(A) = \sum_{k=1}^{n} m(A_k) + m\left(A \setminus \bigcup_{k=1}^{n} A_k\right).$$

Wegen $m\left(A \setminus \bigcup_{k=1}^{n} A_k\right) \geqq 0$ folgt aus dieser Gleichheit die Aussage 1. Weiter ist für beliebige $A_1, A_2 \in \mathfrak{R}_m$

$$m(A_1 \cup A_2) = m(A_1) + m(A_2) - m(A_1 \cap A_2) \leqq m(A_1) + m(A_2).$$

Daraus folgt durch vollständige Induktion

$$m\left(\bigcup_{k=1}^{n} A_k\right) \leqq \sum_{k=1}^{n} m(A_k).$$

Für $A \subset \bigcup_{k=1}^{n} A_k$ ergibt sich aus der Additivität des Maßes

$$m(A) = m\left(\bigcup_{k=1}^{n} A_k\right) - m\left(\bigcup_{k=1}^{n} A_k \setminus A\right) \leqq m\left(\bigcup_{k=1}^{n} A_k\right),$$

woraus zusammen mit der vorhergehenden Ungleichung die Aussage 2 folgt.

Wir haben die Aussagen 1 und 2 für Maße bewiesen, die auf einem Ring gegeben sind. Da aber jedes Maß, das auf einem Semiring gegeben ist, auf einen Ring (unter Beibehaltung seiner Werte auf dem Halbring) fortgesetzt werden kann, sind die Aussagen 1 und 2 auch für Maße richtig, die nur auf einem Semiring gegeben sind.

5.2.3. Die σ-Additivität. Für viele Fragen der Analysis ist es notwendig, Vereinigungen nicht nur endlich vieler Mengen, sondern auch abzählbar vieler Mengen zu betrachten. Im Zusammenhang damit wird die Forderung der Additivität des Maßes (vgl. Definition 1) durch die stärkere Forderung der *σ-Additivität* ersetzt.

Definition 3. Ein Maß m heißt *abzählbar additiv* oder *σ-additiv*, wenn für beliebige Mengen $A, A_1, A_2, \ldots, A_n, \ldots$ aus seinem Definitionsgebiet $\mathfrak{S}_m$ mit den Eigenschaften

$$A = \bigcup_{n=1}^{\infty} A_n, \qquad A_i \cap A_j = \emptyset \quad \text{für} \quad i \neq j,$$

die Beziehung

$$m(A) = \sum_{n=1}^{\infty} m(A_n)$$

gilt.

Das ebene Lebesguesche Maß, das wir in 5.1. konstruiert haben, ist σ-additiv (Satz 8 aus 5.1.). Zu einem Beispiel für ein σ-additives Maß ganz anderer Art kommt man auf folgende Weise. Es sei $X = \{x_1, x_2, \ldots\}$ eine beliebige abzählbare Menge und $\{p_n\}$ eine Folge von positiven Zahlen mit der Eigenschaft $\sum_{n=1}^{\infty} p_n = 1$. Definieren wir für eine beliebige Untermenge A von X

$$m(A) = \sum_{x_n \in A} p_n,$$

so ergibt sich leicht, daß $m(A)$ ein σ-additives Maß mit $m(X) = 1$ auf der σ-Algebra aller Untermengen von X ist. Dieses Maß tritt im Zusammenhang mit vielen Fragen der Wahrscheinlichkeitsrechnung auf.

Ein Beispiel für ein additives, aber nicht σ-additives Maß erhält man durch folgende Konstruktion. Es sei X die Menge aller rationalen Punkte des Intervalls $[0, 1]$ und $\mathfrak{S}_m$ das Mengensystem, das aus allen Durchschnitten der Menge X mit beliebigen in $[0, 1]$ gelegenen Intervallen (a, b), Segmenten $[a, b]$ und Halbsegmenten $[a, b)$, $(a, b]$ besteht. Man bestätigt leicht, daß $\mathfrak{S}_m$ ein Semiring ist. Für die Menge $A_{ab} \in \mathfrak{S}_m$, die durch die Zahlen a, b charakterisiert wird, setzen wir

$$m(A_{ab}) = b - a.$$

Dieses Maß ist additiv, aber nicht σ-additiv. Denn auf Grund von $m(X) = 1$ und der Darstellungsmöglichkeit von X als Vereinigung abzählbar vieler einpunktiger Mengen vom Maß 0 würde die σ-Additivität von m zu einem Widerspruch führen.

In diesem und im nächsten Abschnitt werden wir nur σ-additive Maße betrachten.

Satz 3. *Wenn ein Maß m auf einem Semiring $\mathfrak{S}_m$ σ-additiv ist, dann ist auch seine Fortsetzung μ auf den Ring $\mathfrak{R}(\mathfrak{S}_m)$ σ-additiv.*

Beweis. Sind A und B_n $(n = 1, 2, \ldots)$ Mengen aus $\mathfrak{R}(\mathfrak{S}_m)$, für die

$$A = \bigcup_{n=1}^{\infty} B_n \quad \text{und} \quad B_s \cap B_r = \emptyset \quad \text{für} \quad s \neq r$$

richtig sei, so gibt es nach Satz 3 aus 1.5. endlich viele Mengen A_j und für jedes n endlich viele Mengen B_{ni} aus $\mathfrak{S}_m$ mit der Eigenschaft

$$A = \bigcup_j A_j, \qquad B_n = \bigcup_i B_{ni}, \qquad n = 1, 2, \ldots,$$

wobei die Mengen der rechten Seiten der Gleichungen jeweils paarweise disjunkt sind.

Ist $C_{nij} = B_{ni} \cap A_j$, so bestätigt man leicht, daß auch die Mengen C_{nij} paarweise disjunkt sind und die Eigenschaften

$$A_j = \bigcup_{n=1}^{\infty} \bigcup_i C_{nij},$$

$$B_{ni} = \bigcup_j C_{nij}$$

haben. Aus der σ-Additivität des Maßes m auf $\mathfrak{S}_m$ folgt nun

$$m(A_j) = \sum_{n=1}^{\infty} \sum_i m(C_{nij}), \tag{3}$$

$$m(B_{ni}) = \sum_j m(C_{nij}) \tag{4}$$

und aus der Definition der Fortsetzung μ auf $\mathfrak{R}(\mathfrak{S}_m)$

$$\mu(A) = \sum_j m(A_j), \tag{5}$$

$$\mu(B_n) = \sum_i m(B_{ni}). \tag{6}$$

Setzen wir (3) in (5) ein und vertauschen die Summationen, was auf Grund der Konvergenz der Reihe möglich ist, so folgt mit (4) und (6)

$$\mu(A) = \sum_{n=1}^{\infty} \mu(B_n),$$

was zu zeigen war.

Da nach diesem Satz die σ-Additivität des Maßes bei seiner Fortsetzung von einem Semiring $\mathfrak{S}_m$ auf den Ring $\mathfrak{R}(\mathfrak{S}_m)$ erhalten bleibt, können wir im weiteren gleich von Anfang an annehmen, daß das Maß auf einem Ring definiert ist.

Der folgende Satz beschreibt analoge Eigenschaften für σ-additive Maße und unendliche Vereinigungen, wie sie in Satz 2 für additive Maße und endliche Vereinigungen angegeben wurden.

Satz 4. *Ist m ein σ-additives Maß auf einem Ring $\mathfrak{R}$ und gehören die Mengen $A, A_1, \ldots, A_n, \ldots$ zu $\mathfrak{R}$, dann folgt*

1_σ. *aus $\bigcup_{k=1}^{\infty} A \subset A$ und $A_i \cap A_j = \emptyset$ für $i \neq j$ die Ungleichung*

$$\sum_{k=1}^{\infty} m(A_k) \leqq m(A);$$

2_σ. *aus $\bigcup_{k=1}^{\infty} A_k \supset A$ die Ungleichung*

$$\sum_{k=1}^{\infty} m(A_k) \geqq m(A)$$

(abzählbare Subadditivität).

Beweis. Wenn die Mengen A_k paarweise disjunkt und alle in A enthalten sind, dann ist nach Satz 2, Eigenschaft 1, für jedes n

$$\sum_{k=1}^{n} m(A_k) \leqq m(A).$$

Betrachten wir diese Ungleichung für $n \to \infty$, so ergibt sich die Aussage 1_σ.

Um die zweite Aussage zu beweisen, bilden wir, ausgehend von $A \subset \bigcup_{k=1}^{\infty} A_k$, die Mengen

$$B_n = (A_n \cap A) \setminus \left(\bigcup_{k=1}^{n-1} A_k\right).$$

Für diese Mengen gilt $B_n \in \mathfrak{R}$, $B_i \cap B_j = \emptyset$ für $i \neq j$ und

$$A = \bigcup_{n=1}^{\infty} B_n, \qquad B_n \subset A_n.$$

Mit der σ-Additivität des Maßes folgt daraus

$$m(A) = \sum_{n=1}^{\infty} m(B_n) \leqq \sum_{n=1}^{\infty} m(A_n),$$

was für die Aussage 2_σ zu zeigen war.

Bemerkung. Wie der Beweis zeigt, ist für die Aussage 1_σ nur die Additivität des Maßes wesentlich. Daher ist die Aussage 1_σ auch für beliebige additive Maße richtig. Beim Nachweis der Aussage 2_σ wurde die σ-Additivität des Maßes benutzt. Das vor Satz 3 angegebene Beispiel eines additiven, aber nicht σ-additiven Maßes zeigt, daß für ein solches Maß die Aussage 2_σ i. a. nicht erfüllt ist. Die dort beschriebene Menge X, $\mu(X) = 1$, wird z. B. durch abzählbar viele (einpunktige) Mengen vom Maß 0 überdeckt.

Die folgenden Überlegungen zeigen, daß die Aussage 2_σ und die σ-Additivität sogar äquivalente Aussagen sind.

Es sei μ ein beliebiges Maß auf einem Semiring $\mathfrak{S}$, und $A, A_1, \ldots, A_n, \ldots$ seien Mengen aus $\mathfrak{S}$ mit $A = \bigcup_{k=1}^{\infty} A_k$, $A_i \cap A_j = \emptyset$ für $i \neq j$. Dann ist auf Grund von Eigenschaft 2_σ (nach dem Vorhergehenden eine Eigenschaft beliebiger Maße)

$$\sum_{k=1}^{\infty} \mu(A_k) \leqq \mu(A).$$

Besitzt nun μ die Eigenschaft 2_σ, so ist (weil A von den Mengen A_k überdeckt wird)

$$\sum_{k=1}^{\infty} \mu(A_k) \geqq \mu(A),$$

insgesamt also

$$\sum_{k=1}^{\infty} \mu(A_k) = \mu(A),$$

d. h., das Maß μ ist σ-additiv.

Auf der Grundlage der Äquivalenz von abzählbarer Subadditivität (Eigenschaft 2_σ) und σ-Additivität eines Maßes kann der Nachweis der σ-Additivität durch den oft einfacheren Nachweis der Eigenschaft 2_σ ersetzt werden.

5.3. Die Lebesguesche Fortsetzung eines Maßes

5.3.1. Die Lebesguesche Fortsetzung eines auf einem Semiring mit Eins definierten Maßes. Ist m ein nicht σ-additives Maß auf einem Semiring $\mathfrak{S}_m$, so werden durch seine Fortsetzung auf den Ring $\mathfrak{R}(\mathfrak{S}_m)$ in hohem Maße bereits alle konstruktiven Möglichkeiten der Ausdehnung eines solchen Maßes auf möglichst umfangreiche Klassen von Mengen ausgeschöpft (vgl. 5.3.4.). Ist dagegen das betrachtete Maß σ-additiv, so kann es, ausgehend von $\mathfrak{S}_m$, auf eine wesentlich umfangreichere Klasse als $\mathfrak{R}(\mathfrak{S}_m)$ fortgesetzt werden. Wir kommen zu dieser in gewissem Sinne sogar maximalen Klasse durch die sogenannte *Lebesguesche Fortsetzung* des Maßes m. In diesem Abschnitt betrachten wir die Lebesguesche Fortsetzung zunächst für ein Maß m, dessen Definitionsgebiet $\mathfrak{S}_m$ eine Eins besitzt. Den allgemeinen Fall untersuchen wir anschließend im nächsten Abschnitt.

Es sei m ein σ-additives Maß auf einem Semiring $\mathfrak{S}_m$ mit der Eins E. Auf dem Mengensystem $\mathfrak{A}$ aller Untermengen von E definieren wir eine Funktion $\mu^*(A)$, die als äußeres Maß bezeichnet wird, wie folgt:

Definition 1. Jeder Menge $A \subset E$ wird die Zahl

$$\mu^*(A) = \inf_{A \subset \bigcup_n B_n} \sum_n m(B_n) \tag{1}$$

als *äußeres Maß* zugeordnet. Dabei wird das Infimum über alle möglichen Überdeckungen der Menge A durch endliche oder abzählbare Systeme von Mengen B_n aus $\mathfrak{S}_m$ gebildet.

Wichtig für alle weiteren Konstruktionen ist die folgende Eigenschaft des äußeren Maßes.

Satz 1 (*Abzählbare Subadditivität*). *Ist $A \subset \bigcup_n A_n$, wobei $\{A_n\}$ ein endliches oder abzählbares Mengensystem ist, dann gilt*

$$\mu^*(A) \leqq \sum_n \mu^*(A_n).$$

Der Beweis dieser Behauptung ergibt sich völlig analog zum Beweis des Satzes 3 aus 5.1.

Definition 2. Die Menge A heißt *Lebesgue-meßbar* (oder *meßbar nach Lebesgue*), wenn sich für jede Zahl $\varepsilon > 0$ eine solche Menge $B \in \mathfrak{R}(\mathfrak{S}_m)$ finden läßt, daß

$$\mu^*(A \triangle B) < \varepsilon$$

ist.

Die Einschränkung der Funktion μ^* auf die meßbaren Mengen heißt *Lebesguesches Maß* (oder einfach *Maß*) und wird mit μ bezeichnet. Offensichtlich sind alle Mengen $A \in \mathfrak{S}_m$ meßbar. Für sie gilt

$$\mu(A) = m(A).$$

Der Nachweis dieser Gleichheit kann genauso wie im Fall ebener Mengen geführt werden (vgl. 5.1.2.).

Aus der Gleichung

$$A \triangle B = (E \setminus A) \triangle (E \setminus B)$$

folgt, daß mit der Menge A stets auch ihr Komplement $E \setminus A$ meßbar ist.

Andere grundlegende Eigenschaften der meßbaren Mengen und des für sie definierten Lebesgueschen Maßes beschreiben die folgenden Sätze.

Satz 2. *Das System $\mathfrak{M}$ aller Lebesgue-meßbaren Mengen ist ein Ring.*

Beweis. Da stets

$$A_1 \cap A_2 = A_1 \setminus (A_1 \setminus A_2)$$

und

$$A_1 \cup A_2 = E \setminus (E \setminus A_1) \cap (E \setminus A_2)$$

ist, genügt es zu zeigen, daß für $A_1 \in \mathfrak{M}$, $A_2 \in \mathfrak{M}$ stets auch $A = (A_1 \setminus A_2) \in \mathfrak{M}$ ist.

Es seien A_1 und A_2 meßbar. Dann gibt es für jedes $\varepsilon > 0$ Mengen $B_1 \in \mathfrak{R}(\mathfrak{S}_m)$ und $B_2 \in \mathfrak{R}(\mathfrak{S}_m)$, so daß

$$\mu^*(A_1 \triangle B_1) < \frac{\varepsilon}{2} \quad \text{und} \quad \mu^*(A_2 \triangle B_2) < \frac{\varepsilon}{2}$$

ist. Setzen wir nun $B = B_1 \setminus B_2$ und benutzen die Beziehung

$$(A_1 \setminus A_2) \triangle (B_1 \setminus B_2) \subset (A_1 \triangle B_1) \cup (A_2 \triangle B_2),$$

so folgt

$$\mu^*(A \triangle B) < \varepsilon,$$

was zu zeigen war.

Bemerkung. Offensichtlich ist E auch Eins im Ring $\mathfrak{M}$, d. h., $\mathfrak{M}$ ist eine *Mengenalgebra.*

Satz 3. *Die Funktion $\mu(A)$ ist auf $\mathfrak{M}$ additiv.*

Ein Beweis für diesen Satz ergibt sich durch wörtliche Wiederholung des Beweises von Satz 6 aus 5.1.

Satz 4. *Die Funktion $\mu(A)$ ist auf $\mathfrak{M}$ σ-additiv.*

Beweis.[1]) Es sei $A = \bigcup_{n=1}^{\infty} A_n$, $A, A_1, A_2, \ldots \in \mathfrak{M}$, $A_i \cap A_j = \emptyset$ für $i \neq j$. Nach Satz 1 ist dann

$$\mu(A) \leqq \sum_n \mu(A_n) \tag{2}$$

[1]) Anm. d. Übers.: Der Beweis folgt auch aus Satz 1 und den Betrachtungen am Schluß von 5.2.3.

und nach Satz 3

$$\mu(A) \geqq \mu\left(\bigcup_{n=1}^{N} A_n\right) = \sum_{n=1}^{N} \mu(A_n)$$

für jedes N, d. h., es gilt auch

$$\mu(A) \geqq \sum_{n=1}^{\infty} \mu(A_n). \tag{3}$$

Aus (2) und (3) folgt nun die Aussage des Satzes.

Bei der Untersuchung des Lebesgueschen Maßes in der Ebene, hatten wir festgestellt, daß nicht nur die endlichen, sondern auch die abzählbaren Vereinigungen und Durchschnitte meßbarer Mengen wieder meßbar sind. Das gilt auch für den allgemeinen Fall, wie der folgende Satz zeigt.

Satz 5. *Das System $\mathfrak{M}$ aller Lebesgue-meßbaren Mengen ist eine σ-Algebra mit der Eins E.*

Beweis. Weil

$$\bigcap_n A_n = E \setminus \bigcup_n (E \setminus A_n)$$

und das Komplement einer meßbaren Menge wieder meßbar ist, genügt es zu zeigen, daß für $A_1, A_2, \ldots, A_n, \ldots \in \mathfrak{M}$ stets auch $A = \bigcup_n A_n$ zu $\mathfrak{M}$ gehört. Der Nachweis für diesen Sachverhalt ergibt sich durch wörtliche Wiederholung des Beweises von Satz 7 aus 5.1.

Wie im Fall des Lebesgueschen Maßes in der Ebene folgt auch im allgemeinen Fall aus der σ-Additivität des Maßes seine *Stetigkeit*, d. h., wird das σ-additive Maß μ auf einer fallenden Kette $A_1 \supset A_2 \supset \cdots \supset A_n \supset \cdots$ meßbarer Mengen A_n aus der σ-Algebra $\mathfrak{M}$ betrachtet und ist $A = \bigcap_n A_n$, dann gilt

$$\mu(A) = \lim_{n\to\infty} \mu(A_n),$$

und für eine wachsende Kette $A_1 \subset A_2 \subset \cdots \subset A_n \subset \cdots$ meßbarer Mengen A_n aus der σ-Algebra $\mathfrak{M}$ und $A = \bigcup_n A_n$ gilt analog

$$\mu(A) = \lim_{n\to\infty} \mu(A_n).$$

Ein Beweis für diese Aussagen ergibt sich durch wörtliche Übertragung der Begründung des entsprechenden Sachverhalts für das ebene Maß (vgl. Satz 9 aus 5.1.).

Da, wie wir festgestellt haben, das System $\mathfrak{M}$ eine σ-Algebra ist und die auf ihm definierte Funktion $\mu(A)$ alle Eigenschaften eines σ-additiven Maßes besitzt, ist folgende Definition gerechtfertigt.

Definition 3. Die Einschränkung μ des (vom Maß m erzeugten) äußeren Maßes μ^* auf das System $\mathfrak{M}$ der meßbaren Mengen heißt *Lebesguesche Fortsetzung* $L(m) = \mu$ des Maßes m.

5.3.2. Die Fortsetzung eines auf einem Semiring ohne Eins definierten Maßes. Besitzt der Semiring $\mathfrak{S}_m$, auf dem das Ausgangsmaß m definiert ist, keine Eins, dann muß die im vergangenen Abschnitt durchgeführte Konstruktion der Lebesgueschen Fortsetzung des Maßes m in einigen Punkten geändert werden. Diese Änderungen sind nicht besonders tiefgreifend.

Wird für das Maß m auf einem Semiring ohne Eins die Definition 1 des äußeren Maßes mit der Einschränkung übernommen, daß das äußere Maß jetzt nur auf einem Mengensystem S_{μ^*} definiert wird, für dessen Mengen A mindestens eine Überdeckung $\bigcup_n B_n$, $B_n \in \mathfrak{S}_m$, mit endlicher Summe $\sum_n m(B_n)$ existiert, und die Definition der Meßbarkeit ohne jede Änderung beibehalten, so bleiben die Sätze 2 bis 4 und die abschließende Definition richtig. Der Beweis dieser Sätze im jetzt betrachteten allgemeinen Fall ergibt sich sofort aus den Beweisen der Sätze 2 bis 4, wenn dort die Schlußfolgerung „mit $A_1 \in \mathfrak{M}$ und $A_2 \in \mathfrak{M}$ ist auch $A_1 \cup A_2 \in \mathfrak{M}$“, die einzige, die von der Existenz der Eins in $\mathfrak{S}_m$ Gebrauch machte, mit Hilfe der offensichtlichen Inklusion

$$(A_1 \cup A_2) \triangle (B_1 \triangle B_2) \subset (A_1 \triangle B_1) \cup (A_2 \triangle B_2)$$

begründet wird.

Der Satz 5 muß im Fall des Semirings ohne Eins durch den folgenden ersetzt werden.

Satz 6. *Für ein beliebiges Ausgangsmaß m ist das System $\mathfrak{M}$ Lebesgue-meßbarer Mengen ein δ-Ring. Die Menge $A = \bigcup_{n=1}^{\infty} A_n$, $A_n \in \mathfrak{M}$ $(n = 1, 2, \ldots)$, ist genau dann meßbar, wenn es für die Zahlen $\mu\left(\bigcup_{n=1}^{N} A_n\right)$ eine von N unabhängige obere Schranke gibt.*

Den Beweis für diese Behauptung überlassen wir dem Leser.

Bemerkung. Die Notwendigkeit der letzten Bedingung in Satz 6 ergibt sich zwangsläufig aus den bisherigen Betrachtungen, die jeweils nur Maße mit endlichen Werten zuließen.

Aus Satz 6 ergibt sich die

Folgerung. *Das System $\mathfrak{M}_A$ aller Mengen $B \in \mathfrak{M}$, die in einer festen Menge $A \in \mathfrak{M}$ enthalten sind, ist eine σ-Algebra.*

So ist zum Beispiel das System aller (bezüglich des üblichen Lebesgueschen Maßes auf der Geraden) meßbaren Untermengen eines beliebigen Intervalls $[a, b]$ eine σ-Algebra.

Abschließend weisen wir noch auf eine Eigenschaft der Lebesgueschen Maße hin. Dazu geben wir folgende

Definition 4. Ein Maß μ heißt *vollständig*, wenn aus $\mu(A) = 0$ und $A' \subset A$ stets die Meßbarkeit der Menge A' folgt.

Offensichtlich ist dann $\mu(A') = 0$.

Ist nun μ die Lebesguesche Fortsetzung eines beliebigen Maßes, so folgt aus $A' \subset A$ und $\mu(A) = 0$ sofort $\mu^*(A') = 0$. Da andererseits jede Menge C mit $\mu^*(C) = 0$ wegen $\emptyset \in \mathfrak{R}(\mathfrak{S}_m)$ und

$$\mu^*(C \triangle \emptyset) = \mu^*(C) = 0$$

Lebesgue-meßbar ist, ergibt sich die Aussage: *Die Lebesguesche Fortsetzung eines beliebigen Maßes ist stets vollständig.*

Ganz allgemein kann festgestellt werden, daß jedes σ-additive Maß auf einer σ-Algebra zu einem vollständigen Maß fortgesetzt werden kann. Dazu ist das Maß, wie man leicht sieht, lediglich auf allen Untermengen einer beliebigen Menge vom Maß Null gleich Null zu setzen.

Ergänzende Bemerkungen

1. Die Voraussetzung, daß das Ausgangsmaß m auf einem Semiring gegeben sei (und nicht auf einem beliebigen Mengensystem), ist für die Eindeutigkeit seiner Fortsetzung wesentlich. Betrachten wir im Einheitsquadrat das System aller vertikalen und horizontalen Rechtecke, deren Höhe bzw. Breite immer gleich 1 ist (vgl. Abb. 18), und ordnen diesen Rechtecken ihren Flächeninhalt als Maß zu, so ist die Fortsetzung dieses Maßes auf die von diesen Rechtecken erzeugte Algebra (sogar σ-Algebra) nicht eindeutig bestimmt. Man gebe mindestens zwei verschiedene Fortsetzungen an.

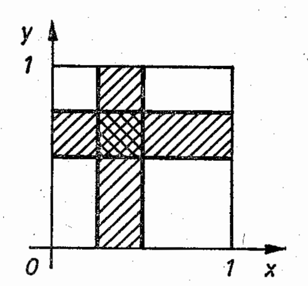


Abb. 18

2. Der Prozeß der Fortsetzung eines Maßes nach LEBESGUE kann auch als Vervollständigungsprozeß eines metrischen Raumes aufgefaßt werden. Führt man auf dem Ring $\mathfrak{R}(\mathfrak{S}_m)$ durch $\varrho(A, B) = m'(A \triangle B)$ eine Metrik ein, so wird $\mathfrak{R}(\mathfrak{S}_m)$ zu einem metrischen (im allgemeinen nicht vollständigen) Raum. Wie man leicht sieht, besteht die Vervollständigung dieses Raumes gerade aus allen Lebesgue-meßbaren Mengen (dabei werden vom metrischen Standpunkt aus Mengen A, B mit $\mu(A \triangle B) = 0$ nicht unterschieden).

Aufgaben

1. Es sei ein Maß m auf einem Semiring $\mathfrak{S}_m$ von Teilmengen aus X mit der Eins X gegeben, μ^* sei das m entsprechende äußere Maß. Man zeige, daß in diesem Fall die Menge A genau dann

nach LEBESGUE meßbar ist, wenn sie nach CARATHÉODORY meßbar ist, d. h., wenn für eine beliebige Teilmenge $Z \subset X$

$$\mu^*(Z) = \mu^*(Z \cap A) + \mu^*(Z \setminus A)$$

gilt.

2. Es sei ein σ-additives Maß m auf einem Ring $\mathfrak{R}$ mit der Eins X gegeben und $m(X) = 1$. Definieren wir für jede Menge $A \subset X$ neben dem äußeren Maß μ^* noch ein inneres Maß μ_* durch

$$\mu_*(A) = 1 - \mu^*(X \setminus A),$$

so ist leicht zu sehen, daß stets $\mu_*(A) \leqq \mu^*(A)$ gilt. Man zeige, daß

$$\mu_*(A) = \mu^*(A) \tag{*}$$

genau dann gilt, wenn A meßbar (im Sinne der Definition 2) ist.

Wir bemerken, daß die Gleichung (*) oft auch zur Definition der Meßbarkeit einer Menge $A \subset X$ benutzt wird, wenn das Ausgangsmaß m bereits auf einem Ring mit der Eins X gegeben ist.

5.3.3. Erweiterung des Begriffes der Meßbarkeit für σ-endliche Maße. Ist das Ausgangsmaß m auf einem beliebigen Semiring ohne Eins von Untermengen der Menge X gegeben, dann erweist sich die oben angegebene Definition der Meßbarkeit als zu eng. Ist beispielsweise X die Ebene, dann sind Mengen mit einem unendlich großen Flächeninhalt wie die ganze Ebene, ein Streifen, das Äußere eines Kreises usw. nicht meßbar. Wir werden daher versuchen, den Begriff der Meßbarkeit so zu erweitern, daß die Gesamtheit der meßbaren Mengen (wie im Fall des auf einem Semiring mit Eins definierten Ausgangsmaßes) eine σ-Algebra (und nicht nur ein δ-Ring) wird. Dabei beschränken wir uns auf den wichtigsten Fall des sogenannten σ-endlichen Maßes, obwohl eine entsprechende Konstruktion auch im allgemeinen Fall durchgeführt werden kann.

Es sei ein σ-additives Maß m auf einem beliebigen Semiring $\mathfrak{S}_m$ von Untermengen der Menge X gegeben. Wir nennen das Maß m *σ-endlich*, wenn die ganze Menge X als abzählbare (aber nicht als endliche) Vereinigung von Mengen aus $\mathfrak{S}_m$ dargestellt werden kann. Ein Beispiel für ein σ-endliches Maß ist der auf der Menge aller Rechtecke der Ebene definierte Flächeninhalt. Ein einfaches Beispiel für ein nicht σ-endliches Maß erhält man auf folgende Weise. Es sei $f(x)$ eine nichtnegative Funktion auf dem Intervall $[0, 1]$ und μ ein Maß mit $\mu(A) = \sum_i f(x_i)$ für jede endliche Untermenge $A = \{x_1, \ldots, x_n\}$ von $[0, 1]$. Wenn die Menge der Punkte x mit $f(x) \neq 0$ nicht abzählbar ist, dann ist das Maß $\mu(A)$ nicht σ-endlich.

Im weiteren sei m ein σ-additives und σ-endliches Maß auf einem Semiring $\mathfrak{S}_m$ von Untermengen der Menge X und $\{B_i\}$ ein entsprechendes abzählbares System von Mengen aus $\mathfrak{S}_m$ mit der Eigenschaft $\bigcup_{i=1}^{\infty} B_i = X$. Wir können o. B. d. A. voraussetzen, daß die Mengen B_i paarweise disjunkt sind. Denn geht man von den B_i zu den Mengen $\tilde{B}_i = B_i \setminus \bigcap_{k=1}^{i-1} B_k$ aus $\mathfrak{R}(\mathfrak{S}_m)$ über, so erhält man X als Vereinigung paar-

weise disjunkter meßbarer Mengen. Da die $\tilde{B}_i$ ihrerseits eine endliche Zerlegung in $\mathfrak{S}_m$ gestatten, ist X auch als abzählbare Vereinigung paarweise disjunkter Mengen aus $\mathfrak{S}_m$ darstellbar.

Führen wir für das Maß m den Lebesgueschen Fortsetzungsprozeß durch, so erhalten wir das Maß μ auf dem δ-Ring $\mathfrak{M}$. Nach der Folgerung aus Satz 6 ist dann für eine beliebige Menge $B \in \mathfrak{M}$ die Gesamtheit

$$M_B = \{C : C \in \mathfrak{M}, C \subset B\}$$

eine σ-Algebra mit der Eins B.

Bezeichnet $\mathfrak{A}$ die Gesamtheit aller Untermengen A von X, für die der Durchschnitt $A \cap B_i$ mit jeder Menge des Systems $\{B_i\}$ meßbar ist, so gilt (was der Leser nachprüfen möge): $\mathfrak{A}$ ist eine σ-Algebra. Ausgehend von der Charakterisierung der Mengen aus $\mathfrak{A}$ ($A \in \mathfrak{A}$ genau dann, wenn $A \cap B_i \in \mathfrak{M}_{B_i}$, $i = 1, 2, \ldots$) und der damit äquivalenten Formulierung ($A \in \mathfrak{A}$ genau dann, wenn es für A eine Darstellung der Form

$$A = \bigcup_{i=1}^{\infty} A_i, \quad A_i \in \mathfrak{M}_{B_i}, \tag{4}$$

gibt) nennt man $\mathfrak{A}$ die *direkte Summe* der σ-Algebren $\mathfrak{M}_{B_i}$.

Die Mengen der σ-Algebra $\mathfrak{A}$ nennen wir *meßbar* (im allgemeinen Sinne) und definieren für sie ein Maß $\tilde{\mu}$ auf folgende Weise: Besitzt A die Darstellung

$$A = \bigcup_{i=1}^{\infty} A_i, \quad A_i \in \mathfrak{M}_{B_i},$$

dann sei

$$\tilde{\mu}(A) = \sum_{i=1}^{\infty} \mu(A_i).$$

Da das Maß μ nichtnegativ ist, strebt die Reihe auf der rechten Seite entweder gegen eine nichtnegative Zahl oder gegen $+\infty$.

Satz 7. *Unter den oben angegebenen Voraussetzungen sind folgende Aussagen richtig:*

1. *Die σ-Algebra $\mathfrak{A}$ und das Maß $\tilde{\mu}$ hängen nicht von der Wahl des Systems $\{B_i\}$ ab; es kann sogar ein beliebiges System $\{B_j^*\}$ von Mengen aus $\mathfrak{M}$ mit $\bigcup_{j=1}^{\infty} B_j^* = X$, $B_l^* \cap B_k^* = \emptyset$ für $l \neq k$, zur Konstruktion von $\mathfrak{A}$ und $\tilde{\mu}$ benutzt werden;*

2. *das Maß $\tilde{\mu}$ ist σ-additiv auf $\mathfrak{A}$;*

3. *die Gesamtheit der Mengen $A \in \mathfrak{A}$ mit $\tilde{\mu}(A) < \infty$ stimmt mit dem δ-Ring $\mathfrak{M}$ überein, und auf diesem Ring ist $\tilde{\mu} = \mu$.*

Beweis. 1. Wir zeigen zunächst, daß $A \in \mathfrak{A}$ genau dann gilt, wenn $A \cap C \in \mathfrak{M}$ für eine beliebige Menge $C \in \mathfrak{M}$ erfüllt ist. Die Hinlänglichkeit dieser Bedingung ist klar; man wählt für C die Mengen $B_i \in \mathfrak{M}$ $(i = 1, 2, \ldots)$. Ihre Notwendigkeit ergibt

sich folgendermaßen. Es sei $A \in \mathfrak{A}$, $C \in \mathfrak{M}$, $\{B_i\}$ das vorgegebene System und ferner $C_i = C \cap B_i$. Dann gilt

$$A \cap C = \bigcup_{i=1}^{\infty} (A \cap C_i), \qquad A \cap C_i \in \mathfrak{M}_{B_i},$$

und

$$\mu\left(\bigcup_{i=1}^{N} (A \cap C_i)\right) \leqq \mu\left(\bigcup_{i=1}^{N} C_i\right) \leqq \mu(C)$$

für jedes N. Aus beiden Beziehungen folgt nach Satz 6 die Meßbarkeit von $A \cap C$.

Ist $\{B_i\}$ das vorgegebene System und $\{B_j^*\}$, $B_j^* \in \mathfrak{M}$, ein beliebiges System paarweise disjunkter Mengen mit $\bigcup_{j=1}^{\infty} B_j^* = \bigcup_{i=1}^{\infty} B_i = X$, so gehören für eine beliebige Menge $A \in \mathfrak{A}$ die Mengen $A \cap B_j^*$ $(j = 1, 2, \ldots)$ nach den Vorbetrachtungen zu $\mathfrak{M}$, d. h., es existiert eine Darstellung

$$A = \bigcup_{j=1}^{\infty} A_j^*, \qquad A_j^* \in \mathfrak{M}_{B_j^*},$$

und das bezüglich $\{B_j^*\}$ berechnete Maß $\tilde{\mu}$ von A stimmt wegen

$$\sum_{j=1}^{\infty} \mu(A_j^*) = \sum_{j=1}^{\infty} \mu(A \cap B_j^*) = \sum_{i,j=1}^{\infty} \mu(A \cap B_j^* \cap B_i) = \sum_{i=1}^{\infty} \mu(A \cap B_i) = \sum_{i=1}^{\infty} \mu(A_i)$$

mit dem für A bezüglich $\{B_i\}$ festgelegten Maß $\tilde{\mu}$ überein. Damit ist die Aussage 1 bewiesen.

2. Es sei $A^{(1)}, A^{(2)}, \ldots \in \mathfrak{A}$, $A^{(k)} \cap A^{(l)} = \emptyset$ für $k \neq l$ und $A = \bigcup_k A^{(k)}$. Aus der σ-Additivität des Maßes μ auf $\mathfrak{M}$ folgt

$$\tilde{\mu}(A) = \sum_{i=1}^{\infty} \mu(A \cap B_i) = \sum_{i,k=1}^{\infty} \mu(A^{(k)} \cap B_i)$$
$$= \sum_{k=1}^{\infty} \left(\sum_{i=1}^{\infty} \mu(A^{(k)} \cap B_i)\right) = \sum_{k=1}^{\infty} \tilde{\mu}(A^{(k)}),$$

d. h., $\tilde{\mu}$ ist σ-additiv.

3. Die letzte Aussage folgt unmittelbar aus Satz 6.

Bemerkung. Die oben beschriebene Erweiterung des Meßbarkeitsbegriffes, die unendliche Werte für das Maß zuläßt, ist auch ohne die Voraussetzung der σ-Endlichkeit des Ausgangsmaßes möglich, z. B. nach folgendem Schema.

Es sei m ein σ-additives Maß in der Menge X, das wir uns schon (nach 5.3.2.) zu einem σ-additiven Maß μ auf einem δ-Ring $\mathfrak{M}$ von Untermengen der Menge X fortgesetzt denken. Ausgehend von $\mathfrak{M}$ nennen wir eine beliebige Menge $A \subset X$ *meßbar bezüglich* $\mathfrak{M}$, wenn $A \cap B \in \mathfrak{M}$ für jede Menge $B \in \mathfrak{M}$ ist. Man verifiziert leicht, daß das System $\mathfrak{A}$ aller bezüglich $\mathfrak{M}$ meßbaren Untermengen von X eine σ-Algebra mit der Eins X ist und daß im Fall einer σ-Algebra $\mathfrak{M}$ $\mathfrak{A} = \mathfrak{M}$ gilt.

Heißt eine Menge $A \in \mathfrak{A}$ eine Nullmenge, wenn $\mu(A \cap B) = 0$ für jede Menge $B \in \mathfrak{M}$ ist, so können wir auf folgende Weise ein Maß $\tilde{\mu}(A)$ auf $\mathfrak{A}$ einführen: Wenn für die gegebene Menge

$A \in \mathfrak{A}$ eine Menge $B \in \mathfrak{M}$ existiert, so daß $A \triangle B$ eine Nullmenge ist, dann sei

$$\tilde{\mu}(A) = \mu(B);$$

existiert keine solche Menge B, dann sei

$$\tilde{\mu}(A) = \infty.$$

Es ist nicht schwierig zu zeigen, daß das Maß $\tilde{\mu}$ σ-additiv ist und auf dem δ-Ring $\mathfrak{M} \subset \mathfrak{A}$ mit μ übereinstimmt.

5.3.4. Die Jordansche Fortsetzung eines Maßes. In 5.2. wurden Maße betrachtet, die nur die Eigenschaft der Additivität besaßen. Jedes solche Maß konnte von einem Semiring $\mathfrak{S}_m$ in eindeutiger Weise auf den minimalen Ring $\mathfrak{R}(\mathfrak{S}_m)$ fortgesetzt werden. Auch eine weitere Ausdehnung auf einen umfassenderen Ring ist möglich (vgl. die Bemerkungen am Anfang von 5.3.1.). Die entsprechende Konstruktion wird als *Fortsetzung des Maßes nach Jordan*[1]) bezeichnet. Ihre Idee, die in Spezialfällen schon von den Mathematikern im alten Griechenland angewandt wurde, besteht in der zweiseitigen Approximation der zu „messenden" Menge A durch Mengen A' und A'' bekannten Maßes mit $A' \subset A \subset A''$ (Approximation von innen und von außen).

Es sei m ein Maß, das auf einem beliebigen Ring $\mathfrak{R}$ gegeben ist.

Definition 5. Eine Menge A heißt *Jordan-meßbar*, wenn für jedes $\varepsilon > 0$ Mengen A' und A'' aus $\mathfrak{R}$ mit

$$A' \subset A \subset A'', \quad m(A'' \setminus A') < \varepsilon,$$

existieren.

Für diese Mengen gilt folgender

Satz 8. *Das System $\mathfrak{R}^*$ aller Jordan-meßbaren Mengen ist ein Ring.*

Es sei $\mathfrak{A}$ das System aller Mengen A, für die es mindestens zwei Mengen B und C aus $\mathfrak{R}$ mit $B \subset A \subset C$ gibt, und es sei für $A \in \mathfrak{A}$

$$\bar{\mu}(A) = \inf_{C \supset A} m(C) \quad (\textit{äußeres Jordansches Maß}),$$

$$\underline{\mu}(A) = \sup_{B \subset A} m(B) \quad (\textit{inneres Jordansches Maß}).$$

Dann ist offensichtlich immer

$$\underline{\mu}(A) \leqq \bar{\mu}(A),$$

und es gelten die folgenden Sätze.

Satz 9. *Der Ring $\mathfrak{R}^*$ stimmt mit dem System aller Mengen A aus $\mathfrak{A}$ überein, für die $\underline{\mu}(A) = \bar{\mu}(A)$ ist.*

Satz 10. *Für $A, A_1, \ldots, A_n \in \mathfrak{A}$, $A \subset \bigcup_{k=1}^{n} A_k$ ist*

$$\bar{\mu}(A) \leqq \sum_{k=1}^{n} \bar{\mu}(A_k).$$

Satz 11. *Für $A, A_1, \ldots, A_n \in \mathfrak{A}$, $A_k \subset A$ $(k = 1, 2, \ldots, n)$ und $A_i \cap A_j = \emptyset$ $(i \neq j)$ ist*

$$\underline{\mu}(A) \geqq \sum_{k=1}^{n} \underline{\mu}(A_k).$$

[1]) Camille Jordan (1838—1922), französischer Mathematiker.

Definieren wir nun, ausgehend von Satz 9, auf $\mathfrak{S}_\mu = \mathfrak{R}^*$ die Funktion μ durch

$$\mu(A) = \underline{\mu}(A) = \bar{\mu}(A),$$

so folgt aus den Sätzen 10, 11 und der offensichtlichen Tatsache, daß für $A \in \mathfrak{R}$

$$\bar{\mu}(A) = \underline{\mu}(A) = m(A)$$

ist, der folgende Satz.

Satz 12. *Die Funktion $\mu(A)$ ist ein Maß und Fortsetzung des Maßes m.*

Die oben dargestellte Konstruktion ist auf ein beliebiges Maß m, das auf einem Ring gegeben ist, anwendbar, insbesondere also auf Mengen in der Ebene. Wird dabei als Ausgangsring die Gesamtheit der Elementarmengen (d. h. der endlichen Vereinigungen von Rechtecken) benutzt, so hängt dieser Ring offensichtlich von der Wahl des Koordinatensystems in der Ebene ab, da Rechtecke mit achsenparallelen Seiten betrachtet werden. Diese Abhängigkeit verliert sich beim Übergang zum ebenen Jordanschen Maß. Wird z. B. neben dem Ring $\mathfrak{R}$ von Elementarmengen bezüglich des Koordinatensystems $\{x_1, x_2\}$ noch ein zweiter Ring $\overline{\mathfrak{R}}$ von Mengen betrachtet, die bezüglich des Koordinatensystems $\{\bar{x}_1, \bar{x}_2\}$ in endlich viele Rechtecke zerlegbar sind, wobei $\{x_1, x_2\}$ und $\{\bar{x}_1, \bar{x}_2\}$ durch eine orthogonale Transformation

$$\bar{x}_1 = x_1 \cdot \cos\alpha + x_2 \cdot \sin\alpha + a_1,$$

$$\bar{x}_2 = -x_1 \cdot \sin\alpha + x_2 \cdot \cos\alpha + a_2$$

verknüpft sein sollen, so erhalten wir in beiden Fällen dasselbe System Jordan-meßbarer Mengen und dasselbe Jordansche Maß μ. Diese Tatsache ergibt sich aus dem folgenden allgemeinen Satz.

Satz 13. *Sind zwei Maße m_1 und m_2 auf den Ringen $\mathfrak{R}_1$ bzw. $\mathfrak{R}_2$ vorgegeben, dann stimmen ihre Jordanschen Fortsetzungen $\mu_1 = j(m_1)$ und $\mu_2 = j(m_2)$ genau dann überein, wenn*

$$\mathfrak{R}_1 \subset \mathfrak{S}_{\mu_2}, \quad m_1(A) = \mu_2(A) \text{ auf } \mathfrak{R}_1,$$

$$\mathfrak{R}_2 \subset \mathfrak{S}_{\mu_1}, \quad m_2(A) = \mu_1(A) \text{ auf } \mathfrak{R}_2$$

gilt.

Für ein Maß m, das nur auf einem Semiring $\mathfrak{S}_m$ gegeben ist, bezeichnen wir natürlicherweise als Jordansche Fortsetzung das Maß

$$j(m) = j(r(m)),$$

das wir im Ergebnis der Fortsetzung von m auf den Ring $\mathfrak{R}(\mathfrak{S}_m)$ und der weiteren Fortsetzung dieses Maßes nach Jordan erhalten.

5.3.5. Eindeutigkeit der Fortsetzung eines Maßes. Wenn eine Menge A Jordan-meßbar bezüglich des Maßes m ist, d. h. zu $\mathfrak{R}^* = \mathfrak{R}^*(\mathfrak{S}_m)$ gehört, dann stimmt der Wert $\tilde{\mu}(A)$ einer beliebigen Fortsetzung des Maßes m auf $\mathfrak{R}^*$ mit dem Wert $J(A)$ der Jordanschen Fortsetzung $J = j(m)$ überein. Man kann weiter zeigen, daß eine Fortsetzung des Maßes m auf ein noch umfassenderes System als die Jordan-meßbaren Mengen $\mathfrak{R}^*$ nicht eindeutig bestimmt ist. Wir präzisieren diesen Sachverhalt im nächsten Satz und legen dazu fest: Eine Menge A heißt *Eindeutigkeitsmenge* für das Maß m, wenn die beiden folgenden Bedingungen erfüllt sind:

1. Es existiert eine Fortsetzung des Maßes m, die für A definiert ist,
2. für zwei beliebige solche Fortsetzungen μ_1 und μ_2 ist

$$\mu_1(A) = \mu_2(A).$$

Satz 14. *Das System der Eindeutigkeitsmengen für das Maß m stimmt mit dem System der Jordan-meßbaren Mengen bezüglich des Maßes m, d. h. mit dem Ring $\mathfrak{R}^*$ überein.*

Wenn wir jedoch nur σ-additive Maße und ihre σ-additiven Fortsetzungen betrachten, ist das System der Eindeutigkeitsmengen im allgemeinen umfassender als $\mathfrak{R}^*$. Für diesen sehr wichtigen Fall der σ-additiven Maße geben wir noch folgende Definition.

Definition 6. Eine Menge A heißt *σ-Eindeutigkeitsmenge* für das σ-additive Maß m, wenn

1. eine σ-additive Fortsetzung λ des Maßes m existiert, die für A definiert ist, d. h. $A \in \mathfrak{S}_\lambda$,
2. für zwei beliebige solche Fortsetzungen λ_1 und λ_2

$$\lambda_1(A) = \lambda_2(A)$$

ist.

Wenn also A eine σ-Eindeutigkeitsmenge für das σ-additive Maß μ ist, dann existiert nach unserer Definition nur ein einziger möglicher Wert $\lambda(A)$ für eine beliebige σ-additive Fortsetzung λ des Maßes μ, falls λ für A definiert ist.

Man sieht leicht, daß jede Jordan-meßbare Menge A auch Lebesgue-meßbar ist und für die Jordan-meßbaren Mengen Jordansches und Lebesguesches Maß übereinstimmen (die Umkehrung gilt nicht; man gebe ein Beispiel dafür an!). Daraus folgt sofort, daß die Jordansche Fortsetzung eines σ-additiven Maßes ebenfalls σ-additiv ist.

Jede Lebesgue-meßbare Menge A ist eine σ-Eindeutigkeitsmenge für das Ausgangsmaß m. Denn nach der Meßbarkeitsdefinition gibt es für jedes $\varepsilon > 0$ eine Menge $B \in \mathfrak{R}$ mit $\mu^*(A \bigtriangleup B) < \varepsilon$. Daraus folgt, daß für eine beliebige σ-additive Fortsetzung λ (die auch A erfaßt)

$$\lambda(A \bigtriangleup B) \leqq \mu^*(A \bigtriangleup B) < \varepsilon$$

gilt. Weil auf Grund der Eindeutigkeit der Fortsetzung m' von m auf $\mathfrak{R} = \mathfrak{R}(\mathfrak{S}_m)$ stets

$$\lambda(B) = m'(B), \quad B \in \mathfrak{R},$$

ist, folgt daraus

$$|\lambda(A) - m'(B)| < \varepsilon.$$

Das bedeutet für zwei beliebige σ-additive Fortsetzungen λ_1 und λ_2 von m

$$|\lambda_1(A) - \lambda_2(A)| < 2\varepsilon,$$

woraus wegen der Willkür in der Wahl von $\varepsilon > 0$

$$\lambda_1(A) = \lambda_2(A)$$

folgt. Damit ist jede Lebesgue-meßbare Menge A eine σ-Eindeutigkeitsmenge. Man kann auch die Umkehrung zeigen, so daß insgesamt gilt: Das System der Lebesgue-meßbaren Mengen bezüglich des Maßes m stimmt mit dem System der σ-Eindeutigkeitsmengen für das Maß m überein.

Ist m ein σ-additives Maß mit dem Definitionsgebiet $\mathfrak{S}$ und $\mathfrak{M} = L(\mathfrak{S})$ das Definitionsgebiet seiner Lebesgueschen Fortsetzung, dann ist die folgende Eigenschaft der Lebesgue-meßbaren Mengen leicht zu verifizieren: Für einen beliebigen Semiring $\mathfrak{S}_1$ mit $\mathfrak{S} \subset \mathfrak{S}_1 \subset \mathfrak{M}$ ist stets

$$L(\mathfrak{S}_1) = L(\mathfrak{S}).$$

5.4. Meßbare Funktionen

5.4.1. Definition und Eigenschaften meßbarer Funktionen. Es seien X und Y zwei beliebige Mengen, $\mathfrak{S}_X$ und $\mathfrak{S}_Y$ zwei Systeme von Untermengen über den entsprechenden Mengen. Eine abstrakte Funktion $y = f(x)$ mit dem Definitionsgebiet X

und Werten in Y heißt $(\mathfrak{S}_X, \mathfrak{S}_Y)$-meßbar, wenn aus $A \in \mathfrak{S}_Y$ stets folgt, daß $f^{-1}(A) \in \mathfrak{S}_X$ ist.

Wenn wir z. B. für X und Y die Zahlengerade nehmen (d. h. reelle Funktionen einer reellen Variablen betrachten) und für $\mathfrak{S}_X$ und $\mathfrak{S}_Y$ das System aller offenen (oder aller abgeschlossenen) Untermengen aus $\boldsymbol{R}^1$, dann geht die oben formulierte Definition der Meßbarkeit in die Definition der Stetigkeit über. Nehmen wir aber für $\mathfrak{S}_X$ und $\mathfrak{S}_Y$ das System aller Borelmengen, so charakterisiert diese Definition die sogenannten *B-meßbaren* (oder *nach Borel meßbaren*) Funktionen.

Im folgenden wird uns der Begriff der Meßbarkeit hauptsächlich vom Standpunkt der Integrationstheorie interessieren. Eine spezielle, aber wichtige Form dieses Begriffes ist für uns deshalb der Begriff der Meßbarkeit einer Zahlenfunktion, die auf einer Menge X mit einem σ-additiven Maß μ definiert ist. Für $\mathfrak{S}_X$ wird dabei die Gesamtheit aller bezüglich μ meßbaren Untermengen von X und für $\mathfrak{S}_Y$ die Gesamtheit aller B-Mengen der Zahlengeraden genommen. Weil jedes σ-additive Maß auf eine σ-Algebra fortgesetzt werden kann, betrachten wir gleich von Anfang an ein System $\mathfrak{S}_X$, das σ-Algebra ist. Auf diese Weise kommen wir für Zahlenfunktionen zu folgender Definition der Meßbarkeit:

Definition 1. Es sei X eine Menge, $\mathfrak{S}_\mu$ eine σ-Algebra von Untermengen der Menge X und μ ein auf $\mathfrak{S}_\mu$ definiertes σ-additives Maß. Eine reelle Funktion $f(x)$ auf X heißt *μ-meßbar*, wenn für jede Borelmenge A der Zahlengeraden stets

$$f^{-1}(A) \in \mathfrak{S}_\mu$$

ist.

Analog heißt eine komplexwertige Funktion φ auf X *μ-meßbar*, wenn $\varphi^{-1}(A) \in \mathfrak{S}_\mu$ für jede Borelmenge A der komplexen Ebene gilt. Man verifiziert leicht, daß diese Bedingung äquivalent ist mit der gleichzeitigen μ-Meßbarkeit des Real- und des Imaginärteiles von φ (einzeln als reelle Funktionen betrachtet).

Eine auf $\boldsymbol{R}^1$ definierte Zahlenfunktion heißt *Borelfunktion* (oder *B-meßbar*), wenn das Urbild jeder Borelmenge stets wieder eine Borelmenge ist.

Satz 1. *Es seien X, Y, Z beliebige Mengen, $\mathfrak{S}_X$, $\mathfrak{S}_Y$, $\mathfrak{S}_Z$ entsprechende Systeme von Untermengen, $y = f(x)$ eine auf X definierte $(\mathfrak{S}_X, \mathfrak{S}_Y)$-meßbare und $z = g(y)$ eine auf Y definierte $(\mathfrak{S}_Y, \mathfrak{S}_Z)$-meßbare Funktion. Dann ist die Funktion*

$$z = \varphi(x) \equiv g(f(x))$$

$(\mathfrak{S}_X, \mathfrak{S}_Z)$-meßbar.

Kurz gesagt: *Eine meßbare Funktion von einer meßbaren Funktion ist eine meßbare Funktion.*

Beweis. Es sei $A \in \mathfrak{S}_Z$. Dann folgt aus der $(\mathfrak{S}_Y, \mathfrak{S}_Z)$-Meßbarkeit von g

$$g^{-1}(A) = B \in \mathfrak{S}_Y,$$

aus der $(\mathfrak{S}_X, \mathfrak{S}_Y)$-Meßbarkeit von f andererseits

$$f^{-1}(B) \in \mathfrak{S}_X.$$

Insgesamt gilt also

$$f^{-1}(g^{-1}(A)) = \varphi^{-1}(A) \in \mathfrak{S}_X,$$

d. h., die Funktion φ ist $(\mathfrak{S}_X, \mathfrak{S}_Z)$-meßbar.

Folgerung. *Eine Borelfunktion von einer μ-meßbaren Zahlenfunktion ist μ-meßbar. Insbesondere ist eine stetige Funktion von einer μ-meßbaren Funktion stets μ-meßbar.*

Wenn keine Mißverständnisse zu befürchten sind, werden wir im weiteren anstelle von „μ-Meßbarkeit" einfach „Meßbarkeit" schreiben.

Satz 2. *Eine reelle Funktion $f(x)$ ist genau dann meßbar, wenn für jede reelle Zahl c die Menge $\{x: f(x) < c\}$ meßbar ist.*

Beweis. Die Notwendigkeit der Bedingung ist klar, da die Halbgerade $(-\infty, c)$ eine Borelmenge ist. Für den Nachweis ihrer Hinlänglichkeit benutzen wir die Tatsache, daß die vom Mengensystem aller Halbgeraden $(-\infty, c)$ erzeugte σ-Algebra mit der σ-Algebra aller Borelmengen der Geraden übereinstimmt. Nach 1.5.5. folgt daraus, daß das Urbild jeder Borelmenge zu der σ-Algebra gehört, die von den Urbildern der Halbgeraden erzeugt wird. Damit folgt aus der Meßbarkeit der Mengen $\{x: f(x) < c\}$ die Meßbarkeit von f.

Die eben bewiesene Eigenschaft meßbarer Funktionen wird auch oft zur Definition der Meßbarkeit einer Funktion benutzt. So nennt man dann eine Funktion $f(x)$ meßbar, wenn alle Mengen $\{x: f(x) < c\}$ meßbar sind.

5.4.2. Operationen mit meßbaren Funktionen. Wir zeigen zunächst, daß die Gesamtheit aller auf einer Menge X definierten meßbaren Funktionen bezüglich der arithmetischen Operationen abgeschlossen ist.

Satz 3. *Summe, Differenz und Produkt zweier meßbarer Funktionen sind meßbar. Der Quotient zweier meßbarer Funktionen ist ebenfalls meßbar, falls der Nenner nicht verschwindet.*

Beweis. Den Beweis dieses Satzes führen wir schrittweise.

1. Wenn f meßbar ist, dann sind offensichtlich auch die Funktionen kf und $a + f$ für beliebige Konstanten k und a meßbar.

2. Weiter ist für meßbare Funktionen f und g stets auch die Menge

$$\{x: f(x) > g(x)\}$$

meßbar. Es ist nämlich

$$\{x: f(x) > g(x)\} = \bigcup_{k=1}^{\infty} (\{x: f(x) > r_k\} \cap \{x: g(x) < r_k\}),$$

wobei die Vereinigung über alle rationalen Zahlen r_k (in beliebiger Reihenfolge) gebildet wird. Daraus folgt, daß

$$\{x: f(x) > a - g(x)\} = \{x: f(x) + g(x) > a\}$$

meßbar ist, d. h., die Summe meßbarer Funktionen ist meßbar.

3. Aus 1. und 2. folgt die Meßbarkeit der Differenz.

4. Das Produkt meßbarer Funktionen ist meßbar. In der Identität

$$fg = \frac{1}{4}\,[(f+g)^2 - (f-g)^2]$$

ist nämlich der rechtsstehende Ausdruck meßbar. Das ergibt sich aus 1. bis 3. und der Folgerung aus Satz 1, nach der das Quadrat einer meßbaren Funktion meßbar ist.

5. Wenn $f(x)$ meßbar und $f(x) \neq 0$ ist, dann ist $1/f(x)$ meßbar. Für $c > 0$ ist nämlich

$$\left\{x\colon \frac{1}{f(x)} < c\right\} = \left\{x\colon f(x) > \frac{1}{c}\right\} \cup \{x\colon f(x) < 0\},$$

für $c < 0$

$$\left\{x\colon \frac{1}{f(x)} < c\right\} = \left\{x\colon 0 > f(x) > \frac{1}{c}\right\}$$

und für $c = 0$

$$\left\{x\colon \frac{1}{f(x)} < c\right\} = \{x\colon f(x) < c\},$$

d. h., $\{x\colon 1/f(x) < c\}$ ist in jedem Fall durch meßbare Mengen darstellbar, also selbst meßbar. Aus 4. und 5. ergibt sich die Meßbarkeit des Quotienten $f(x)/g(x)$ (unter der Bedingung $g(x) \neq 0$).

Damit haben wir gezeigt, daß arithmetische Operationen, angewandt auf meßbare Funktionen, wieder zu meßbaren Funktionen führen.

Wir zeigen nun, daß die Gesamtheit der meßbaren Funktionen nicht nur bezüglich der arithmetischen Operationen, sondern auch bezüglich der Grenzwertbildung abgeschlossen ist.

Satz 4. *Der Grenzwert einer für jedes $x \in X$ konvergenten Folge meßbarer Funktionen ist meßbar.*

Beweis. Es strebe $f_n(x) \to f(x)$ für jedes $x \in X$, dann ist

$$\{x\colon f(x) < c\} = \bigcup_{k>0} \bigcup_{n} \bigcap_{m>n} \left\{x\colon f_m(x) < c - \frac{1}{k}\right\}. \tag{1}$$

Denn für $x_0 \in \{x\colon f(x) < c\}$ gibt es eine Zahl k_0, so daß auch $f(x_0) < c - 2/k_0$. Andererseits kann man bei vorgegebenem k_0 stets eine Zahl n_0 finden, so daß für alle $m \geqq n_0$

$$f_m(x_0) < c - \frac{1}{k_0}$$

ist, d. h., x_0 gehört zur rechten Seite von (1).

Wird umgekehrt vorausgesetzt, daß x_1 zur rechten Seite von (1) gehört, so ist $x_1 \in \{x: f_m(x) < c - 1/k_1\}$ für ein gewisses k_1 und alle $m \geqq n_1$. Damit ist dann auch $f(x_1) < c$, d. h., x_1 gehört zur linken Seite von (1).

Aus der Meßbarkeit aller Funktionen $f_n(x)$ folgt die Meßbarkeit der Mengen

$$\left\{x: f_m(x) < c - \frac{1}{k}\right\}$$

Da die Gesamtheit aller meßbaren Mengen eine σ-Algebra bildet, ist nach (1) die Menge $\{x: f(x) < c\}$ ebenfalls meßbar.

Wie in 5.3.2. schon bemerkt wurde, kann jedes σ-additive Maß μ, das auf einer σ-Algebra von Untermengen einer Menge X erklärt ist, o. B. d. A. zu einem vollständigen Maß erweitert werden, d. h., man kann annehmen, daß für jede meßbare Menge A mit $\mu(A) = 0$ auch jede ihrer Untermengen $A' \subset A$ meßbar und natürlich $\mu(A') = 0$ ist. Im weiteren werden wir deshalb stets annehmen, daß die betrachteten Maße vollständig sind.

5.4.3. Äquivalenz. Bei der Untersuchung meßbarer Funktionen kann man oft ihre Werte auf einer Menge von Maß Null vernachlässigen. In Zusammenhang damit steht folgende Definition.

Definition 2. Zwei Funktionen f und g, die auf ein und derselben meßbaren Menge E gegeben sind, heißen *äquivalent* (geschrieben: $f \sim g$), wenn

$$\mu\{x: f(x) \neq g(x)\} = 0$$

ist.

Man sagt, daß eine Eigenschaft *fast überall* auf E erfüllt ist, wenn sie überall auf E mit Ausnahme einer Punktmenge vom Maß Null erfüllt ist. Mit dieser Sprechweise heißen also zwei Funktionen äquivalent, wenn sie fast überall gleich sind.

Satz 5. *Ist $f(x)$ eine auf einer meßbaren Menge E definierte Funktion und auf dieser Menge äquivalent zu einer meßbaren Funktion $g(x)$, dann ist auch $f(x)$ meßbar.*

Beweis. Aus der Definition der Äquivalenz folgt, daß die Mengen

$$\{x: f(x) < a\} \quad \text{und} \quad \{x: g(x) < a\}$$

sich voneinander höchstens um eine Menge vom Maß Null unterscheiden. Da das Maß (entsprechend der allgemeinen Voraussetzung) vollständig und die zweite Menge meßbar ist, ergibt sich daraus die Meßbarkeit der ersten Menge.

Bemerkung. In der klassischen Analysis spielt der Begriff der Äquivalenz von Funktionen keine wesentliche Rolle, da dort hauptsächlich stetige Funktionen betrachtet werden, für die die Begriffe Äquivalenz und Gleichheit übereinstimmen. Präziser ausgedrückt, sind f und g zwei auf einem Intervall E definierte stetige Funktionen, die dort (bezüglich des Lebesgueschen Maßes) äquivalent sind, so würde aus ihrer Verschiedenheit in einem Punkt x_0, $f(x_0) \neq g(x_0)$, auf Grund der Stetigkeit

$f(x) \neq g(x)$ für eine gewisse Umgebung von x_0 mit einem positiven Maß folgen, was der Äquivalenz widersprechen würde, d. h., stetige Funktionen können nur äquivalent sein, wenn sie übereinstimmen.

Für beliebige meßbare Funktionen bedeutet die Äquivalenz durchaus nicht ein Übereinstimmen. Beispielsweise ist die Funktion, die in allen rationalen Punkten der Geraden gleich 1 und in allen irrationalen Punkten gleich 0 ist, äquivalent zur Funktion, die identisch 0 ist.

5.4.4. Konvergenz fast überall. Weil in vielen Fällen das Verhalten einer meßbaren Funktion auf einer beliebigen Menge vom Maß Null für uns unwesentlich ist, führen wir folgende natürliche Verallgemeinerung des üblichen Begriffes der punktweisen Konvergenz ein.

Definition 3. Eine Funktionenfolge $\{f_n(x)\}$, die auf einer meßbaren Menge X definiert ist, heißt *fast überall konvergent* (auf X) gegen die Funktion $f(x)$, wenn

$$\lim_{n\to\infty} f_n(x) = f(x) \tag{2}$$

für fast alle $x \in X$ ist (d. h., wenn die Menge aller Punkte, für die (2) nicht erfüllt ist, das Maß Null hat).

Beispiel. Die Funktionenfolge $f_n(x) = (-x)^n$ konvergiert auf dem Intervall $[0, 1]$ für $n \to \infty$ fast überall gegen die Funktion $f(x) \equiv 0$ (überall mit Ausnahme des Punktes $x = 1$).

Der Satz 4 erlaubt nun folgende Verallgemeinerung.

Satz 4′. *Wenn eine Folge meßbarer Funktionen $f_n(x)$ fast überall auf X gegen die Funktion $f(x)$ konvergiert, dann ist $f(x)$ auch auf X meßbar.*

Beweis. Es sei A die Menge, auf der

$$\lim_{n\to\infty} f_n(x) = f(x)$$

gilt. Nach Voraussetzung ist dann $\mu(X \setminus A) = 0$. Die Funktion $f(x)$ ist meßbar auf A nach Satz 4 und meßbar auf $X \setminus A$, weil offensichtlich jede Funktion auf einer Nullmenge meßbar ist, d. h., $f(x)$ ist auf der ganzen Menge X meßbar.

Aufgabe. Es sei $\{f_0(x)\}$ eine Folge meßbarer Funktionen, die fast überall gegen eine Grenzfunktion $f(x)$ konvergiert. Man zeige, daß die Folge $\{f_n(x)\}$ dann und nur dann gegen eine Funktion $g(x)$ fast überall konvergiert, wenn $g(x)$ zu $f(x)$ äquivalent ist (d. h., die Grenzfunktion einer fast überall konvergenten Folge ist bis auf Äquivalenz eindeutig bestimmt).

5.4.5. Der Satz von Egorov. Im Jahre 1911 wurde von D. F. Egorov folgender wichtiger Zusammenhang zwischen der Konvergenz fast überall und der gleichmäßigen Konvergenz entdeckt.

Satz 6. *Es sei* $\{f_n(x)\}$ *eine Folge meßbarer Funktionen, die auf* E *fast überall gegen* $f(x)$ *konvergiert. Dann existiert für jedes* $\delta > 0$ *eine meßbare Menge* $E_\delta \subset E$, *so daß*

1. $\mu(E_\delta) > \mu(E) - \delta$ *ist;*
2. *auf* E_δ *die Folge* $\{f_n(x)\}$ *gleichmäßig gegen* $f(x)$ *konvergiert.*

Beweis. Für die meßbaren Funktionen $f_n(x)$ und $f(x)$ (vgl. Satz 4′) setzen wir

$$E_n{}^m = \bigcap_{i \geqq n} \left\{x\colon |f_i(x) - f(x)| < \frac{1}{m}\right\},$$

d. h., bei festem m und n bezeichnet $E_n{}^m$ die Menge aller $x \in E$ mit

$$|f_i(x) - f(x)| < \frac{1}{m}$$

für alle $i \geqq n$. Aus dieser Definition ergibt sich sofort

$$E_1{}^m \subset E_2{}^m \subset \cdots \subset E_n{}^m \subset \cdots.$$

Schreiben wir

$$\bigcup_{n=1}^{\infty} E_n{}^m = E^m,$$

so folgt aus der Stetigkeit des σ-additiven Maßes: Für jedes m und jedes $\delta > 0$ gibt es eine Zahl $n_0(m)$, so daß

$$\mu(E^m \setminus E^m_{n_0(m)}) < \frac{\delta}{2^m}$$

ist. Setzen wir nun

$$E_\delta = \bigcap_{m=1}^{\infty} E^m_{n_0(m)},$$

so erfüllt diese Menge, wie im folgenden gezeigt wird, die Forderungen des Satzes.

Daß die Folge $\{f_i(x)\}$ gleichmäßig auf E_δ gegen $f(x)$ konvergiert, ergibt sich aus der Ungleichung

$$|f_i(x) - f(x)| < \frac{1}{m} \qquad \text{für } i > n_0(m),$$

die bei einem beliebig vorgegebenen m für jedes $x \in E_\delta$ erfüllt ist. Zur Abschätzung des Maßes der Menge $E \setminus E_\delta$ stellen wir zunächst fest, daß $\mu(E \setminus E^m) = 0$ für jedes m ist. Denn ist $x_0 \in E \setminus E^m$, so gibt es stets beliebig große Indexwerte i mit

$$|f_i(x_0) - f(x_0)| \geqq \frac{1}{m},$$

d. h., $f_n(x)$ konvergiert in x_0 nicht gegen $f(x)$. Weil aber $\{f_n(x)\}$ fast überall gegen $f(x)$

konvergiert, muß

$$\mu(E \setminus E^m) = 0$$

sein. Aus dieser Tatsache folgt

$$\mu(E \setminus E^m_{n_0(m)}) = \mu(E^m \setminus E^m_{n_0(m)}) < \frac{\delta}{2^m}.$$

Damit ergibt sich für das Maß der Menge $E \setminus E_\delta$

$$\mu(E \setminus E_\delta) = \left(E \setminus \bigcap_{m=1}^{\infty} E^m_{n_0(m)}\right) = \mu\left(\bigcup_{m=1}^{\infty} (E \setminus E^m_{n_0(m)})\right)$$
$$\leqq \sum_{m=1}^{\infty} \mu(E \setminus E^m_{n_0(m)}) < \sum_{m=1}^{\infty} \frac{\delta}{2^m} = \delta,$$

was zu zeigen war.

5.4.6. Konvergenz dem Maß nach

Definition 4. Eine Folge $\{f_n(x)\}$ meßbarer Funktionen *konvergiert dem Maß nach* gegen die Funktion $f(x)$, wenn für jedes $\sigma > 0$

$$\lim_{n\to\infty} \mu(\{x: |f_n(x) - f(x)| \geqq \sigma\}) = 0$$

gilt.

Die beiden folgenden Sätze 7 und 8 zeigen den Zusammenhang zwischen der Konvergenz fast überall und der Konvergenz dem Maß nach.

Satz 7. *Konvergiert eine Folge $\{f_n(x)\}$ meßbarer Funktionen fast überall gegen eine Funktion $f(x)$, dann konvergiert sie auch dem Maß nach gegen dieselbe Grenzfunktion $f(x)$.*

Beweis. Es sei A die Menge $(\mu(A) = 0)$, auf der $\{f_n(x)\}$ nicht gegen $f(x)$ strebt,

$$E_k(\sigma) = \{x: |f_k(x) - f(x)| \geqq \sigma\},$$

$$R_n(\sigma) = \bigcup_{k=n}^{\infty} E_k(\sigma),$$

$$M = \bigcap_{n=1}^{\infty} R_n(\sigma).$$

Die Mengen $E_k(\sigma)$ sind auf Grund der Meßbarkeit der Funktionen $f_k(x)$ und $f(x)$ (nach Satz 4') meßbar und damit auch die anderen Mengen. Aus der offensichtlichen Beziehung

$$R_1(\sigma) \supset R_2(\sigma) \supset \cdots$$

ergibt sich mit der Stetigkeit des Maßes

$$\mu\big(R_n(\sigma)\big) \to \mu(M) \quad \text{für } n \to \infty.$$

Wie im folgenden gezeigt wird, ist

$$M \subset A, \quad \text{d. h. } \mu(M) = 0. \tag{3}$$

Damit gilt

$$\mu\big(R_n(\sigma)\big) \to 0 \quad \text{für } n \to \infty$$

und wegen der offensichtlichen Inklusion $E_n(\sigma) \subset R_n(\sigma)$ auch

$$\mu\big(E_n(\sigma)\big) \to 0 \quad \text{für } n \to \infty,$$

also die Aussage des Satzes.

Wir zeigen nun (3). Ist $x_0 \notin A$, dann gilt

$$\lim_{n\to\infty} f_n(x_0) = f(x_0),$$

d. h., für jedes $\sigma > 0$ gibt es ein n_0, so daß

$$|f_n(x_0) - f(x_0)| < \sigma \quad \text{für } n > n_0$$

ist. Das bedeutet $x_0 \notin E_n(\sigma)$ für $n > n_0$ und damit auch $x_0 \notin M$, also $M \subset A$, was zu zeigen war.

Man überzeugt sich leicht davon, daß für eine Funktionenfolge aus der Konvergenz dem Maß nach im allgemeinen nicht ihre Konvergenz fast überall folgt. Wir konstruieren dafür ein Beispiel.

Auf dem halboffenen Intervall (0, 1] definieren wir für jede natürliche Zahl k die Funktionen

$$f_1^{(k)}, f_2^{(k)}, \ldots, f_k^{(k)}$$

durch

$$f_i^{(k)} = \begin{cases} 1 \ \text{für } \dfrac{i-1}{k} < x \leqq \dfrac{i}{k}, \\ 0 \ \text{für die restlichen } x \text{ aus } (0, 1] \end{cases}$$

und numerieren sie in der Reihenfolge ihres Auftretens. Die so gebildete Folge konvergiert dem Maß nach gegen 0, aber konvergiert sonst in keinem einzigen Punkt des Intervalls (0, 1]. (Den Beweis überlassen wir dem Leser.)

Aufgabe. Es sei $\{f_n(x)\}$ eine Folge meßbarer Funktionen, die dem Maß nach gegen eine Grenzfunktion $f(x)$ konvergiert. Man zeige, daß die Folge $\{f_n(x)\}$ dann und nur dann gegen eine Funktion $g(x)$ dem Maß nach konvergiert, wenn $g(x)$ zu $f(x)$ äquivalent ist (d. h., die Grenzfunktion einer dem Maß nach konvergenten Folge ist bis auf Äquivalenz eindeutig bestimmt).

Das oben angegebene Beispiel zeigt, daß sich der Satz 7 nicht vollständig umkehren läßt. Eine etwas abgeschwächte Form der Umkehrung von Satz 7 ist der folgende Satz.

Satz 8. *Es sei $\{f_n(x)\}$ eine Folge meßbarer Funktionen, die dem Maß nach gegen $f(x)$ konvergiert. Dann kann man aus dieser Folge eine Teilfolge $\{f_{n_k}(x)\}$ auswählen, die fast überall gegen $f(x)$ konvergiert.*

Beweis. Für zwei beliebige Folgen positiver Zahlen $\{\varepsilon_n\}$ und $\{\eta_k\}$ mit der Eigenschaft

$$\lim_{n\to\infty} \varepsilon_n = 0 \quad \text{bzw.} \quad \sum_{k=1}^{\infty} \eta_k < \infty$$

konstruieren wir eine Indexfolge

$$n_1 < n_2 < \cdots$$

auf folgende Weise: Wir wählen einen Index n_1 in der Folge $\{f_n(x)\}$ so aus, daß

$$\mu(\{x\colon |f_{n_1}(x) - f(x)| \geqq \varepsilon_1\}) < \eta_1$$

ist (die Existenz eines solchen Index ist klar). Dann wählen wir $n_2 > n_1$ so, daß

$$\mu(\{x\colon |f_{n_2}(x) - f(x)| \geqq \varepsilon_2\}) < \eta_2$$

ist, usw. Allgemein wird $n_k > n_{k-1}$ mit

$$\mu(\{x\colon |f_{n_k}(x) - f(x)| \geqq \varepsilon_k\}) < \eta_k$$

gewählt. Die dieser Indexfolge entsprechende Teilfolge $\{f_{n_k}(x)\}$ erfüllt, wie wir im folgenden zeigen, die Forderungen des Satzes. Ist

$$R_i = \bigcup_{k=i}^{\infty} \{x\colon |f_{n_k}(x) - f(x)| \geqq \varepsilon_k\}, \qquad Q = \bigcap_{i=1}^{\infty} R_i,$$

so folgt aus der Stetigkeit des Maßes für $R_1 \supset R_2 \supset \cdots \supset R_n \supset \cdots$ die Beziehung

$$\lim_{i\to\infty} \mu(R_i) = \mu(Q).$$

Andererseits ist $\mu(R_i) < \sum_{k=i}^{\infty} \eta_k$ und daher $\mu(R_i) \to 0$ für $i \to \infty$, d. h., es ist $\mu(Q) = 0$. Damit bleibt noch zu zeigen, daß die Folge $\{f_{n_k}(x)\}$ in allen Punkten von $E \setminus Q$ gegen $f(x)$ konvergiert. Es sei $x_0 \in E \setminus Q$. Dann existiert ein i_0, so daß $x_0 \notin R_{i_0}$, d. h.

$$x_0 \notin \{x\colon |f_{n_k}(x) - f(x)| \geqq \varepsilon_k\} \qquad \text{für alle } k \geqq i_0$$

oder

$$|f_{n_k}(x_0) - f(x_0)| < \varepsilon_k \qquad \text{für alle } k \geqq i_0.$$

Auf Grund der Voraussetzung $\varepsilon_k \to 0$ folgt daraus

$$\lim_{k\to\infty} f_{n_k}(x_0) = f(x_0),$$

was zu zeigen war.

5.4.7. Der Satz von Lusin. Die C-Eigenschaft. Die Definition der Meßbarkeit einer Funktion, wie sie am Anfang von 5.4. gegeben worden ist, bezieht sich auf Funktionen auf beliebigen Mengen und ist im allgemeinen Fall überhaupt nicht mit dem Begriff der Stetigkeit verknüpft. Werden jedoch Funktionen auf abgeschlossenen Intervallen betrachtet, dann gibt es einen wichtigen Zusammenhang zwischen Meßbarkeit und Stetigkeit, der 1913 von N. N. Lusin entdeckt wurde.

Satz 9. *Ist $f(x)$ eine auf dem abgeschlossenen Intervall $[a, b]$ definierte Funktion, so ist $f(x)$ genau dann meßbar, wenn für jedes $\varepsilon > 0$ eine auf $[a, b]$ stetige Funktion $\varphi_\varepsilon(x)$ existiert, so daß*

$$\mu(\{x\colon f(x) \neq \varphi_\varepsilon(x)\}) < \varepsilon$$

ist.

Eine meßbare Funktion kann also durch Änderung auf einer Menge beliebig kleinen Maßes zu einer stetigen Funktion auf $[a, b]$ gemacht werden. Von einer Funktion auf $[a, b]$, die mit Hilfe einer „beliebig kleinen Deformation" stetig gemacht werden kann, sagt man, daß sie die *C-Eigenschaft* (Bezeichnung von N. N. Lusin) besitzt. Der Satz von Lusin zeigt, daß für eine Funktion einer reellen Variablen die C-Eigenschaft zur Definition der Meßbarkeit verwendet werden kann. Einen Beweis für den Satz von Lusin kann man unter Verwendung des Satzes von Egorov erhalten. (Der Leser führe diesen Beweis selbst durch!)

5.5. Das Lebesguesche Integral

Der Begriff des Riemannschen Integrals, wie er in den Grundvorlesungen zur Analysis entwickelt wird, ist nur für stetige Funktionen und Funktionen mit „nicht allzu vielen" Unstetigkeitspunkten verwendbar. Für meßbare Funktionen, die auf ihrem gesamten Definitionsgebiet unstetig oder auf einer abstrakten Menge (ohne Stetigkeitsbegriff) definiert sein können, ist die Riemannsche Konstruktion des Integrals unbrauchbar. Die Aufgabe der Integration solcher Funktionen wird durch einen vollkommeneren Integralbegriff gelöst, der von Lebesgue eingeführt wurde.

Die grundlegende Idee der Konstruktion des Lebesgueschen Integrals besteht darin, daß im Unterschied zum Riemannschen Integral die Integrationspunkte x nicht auf Grund ihrer Nähe auf der x-Achse, sondern auf Grund der Nähe der Funktionswerte in diesen Punkten zusammengefaßt werden. Dieses Verfahren macht eine Erweiterung des Integralbegriffs auf eine sehr umfangreiche Klasse von Funktionen möglich. Außerdem kann das Lebesguesche Integral völlig einheitlich für Funktionen definiert werden, die auf verschiedensten Mengen mit Maß gegeben sind, während das Riemannsche Integral zuerst für Funktionen einer Variablen eingeführt und dann mit entsprechenden Änderungen auf den Fall mehrerer Variabler übertragen wird. Für den Fall abstrakter Mengen besitzt das Riemannsche Integral im allgemeinen keinen Sinn.

Überall, wo nicht ausdrücklich das Gegenteil vorausgesetzt wird, betrachten wir im folgenden ein vollständiges σ-additives Maß, das auf einer σ-Algebra von Mengen mit der Eins X definiert sei; alle betrachteten Mengen $A \subset X$ werden wir als meßbar voraussetzen; die Funktionen $f(x)$ seien für $x \in X$ definiert und meßbar.

Es ist günstig, das Lebesguesche Integral zuerst für die sogenannten *Treppenfunktionen* zu definieren und es dann auf umfangreichere Funktionenklassen zu erweitern. In 5.5.2. bis 5.5.4. wird die Konstruktion des Lebesgueschen Integrals für den Fall durchgeführt, daß das Maß der ganzen Menge X endlich ist. Der Fall des unendlichen Maßes wird in 5.5.6. betrachtet.

5.5.1. Treppenfunktionen

Definition 1. Eine Funktion $f(x)$, die auf einer Menge X mit Maß definiert sei, heißt *Treppenfunktion*, wenn sie meßbar ist und höchstens abzählbar viele verschiedene Funktionswerte annimmt.

Die Struktur der Treppenfunktionen charakterisiert der folgende Satz.

Satz 1. *Eine Funktion $f(x)$ mit höchstens abzählbar vielen verschiedenen Funktionswerten*

$$y_1, y_2, \ldots, y_n, \ldots$$

ist genau dann meßbar, wenn alle Mengen

$$A_n = \{x: f(x) = y_n\}$$

meßbar sind.

Beweis. Die Notwendigkeit der Bedingung ist klar, da jedes A_n Urbild der einpunktigen Borelmenge $\{y_n\}$ ist. Ihre Hinlänglichkeit folgt aus der Tatsache, daß unter den Voraussetzungen des Satzes das Urbild $f^{-1}(B)$ jeder Borelmenge B als abzählbare Vereinigung $\bigcap\limits_{y_n \in B} A_n$ der meßbaren Mengen A_n dargestellt werden kann und demzufolge meßbar ist.

Die Verwendung der Treppenfunktionen bei der Konstruktion des Lebesgueschen Integrals für beliebige meßbare Funktionen basiert auf folgendem Satz.

Satz 2. *Eine Funktion $f(x)$ ist genau dann meßbar, wenn sie als Grenzfunktion einer gleichmäßig konvergenten Folge von Treppenfunktionen dargestellt werden kann.*

Beweis. Die Hinlänglichkeit der Bedingung ist nach Satz 4 aus 5.4. klar. Um ihre Notwendigkeit zu zeigen, konstruieren wir zu der vorgegebenen meßbaren Funktion $f(x)$ die Funktionen

$$f_n(x) = \frac{m}{n} \quad \text{für} \quad \frac{m}{n} \leqq f(x) < \frac{m+1}{n}, \quad n = 1, 2, \ldots,$$

wobei m jeweils bei festem n die ganzen Zahlen durchläuft. Offensichtlich sind die $f_n(x)$ Treppenfunktionen und konvergieren wegen $|f(x) - f_n(x)| \leqq 1/n$ gleichmäßig gegen $f(x)$.

5.5.2. Das Lebesguesche Integral für Treppenfunktionen. Wir führen zuerst den Begriff des Lebesgueschen Integrals für den einfachen Fall ein, daß die Integranden Treppenfunktionen sind.

Es sei f eine Treppenfunktion, die auf der meßbaren Menge $A \subset X$ die Werte

$$y_1, y_2, \ldots, y_n, \ldots, \qquad y_i \neq y_j \text{ für } i \neq j,$$

annimmt.

In natürlicher Weise definieren wir das Integral der Funktion f über die Menge A durch die Gleichung

$$\int_A f(x)\, d\mu = \sum_n y_n \mu(A_n) \quad \text{mit} \quad A_n = \{x : x \in A, f(x) = y_n\}, \tag{1}$$

wenn die Reihe konvergiert. Damit kommen wir zur folgenden Definition (in der wir aus früher erklärten Gründen die absolute Konvergenz der Reihe fordern).

Definition 2. Eine Treppenfunktion f heißt *integrierbar* oder *summierbar* (bezüglich des Maßes μ) auf der Menge A, wenn die Reihe (1) absolut konvergiert. Ist f integrierbar, dann heißt die Summe der Reihe (1) *Integral* von f über die Menge A.

In dieser Definition wurden alle y_n als verschieden vorausgesetzt. Man kann jedoch den Wert des Integrals einer Treppenfunktion auch als Summe von Produkten der Art $c_k \mu(B_k)$ darstellen, wobei nicht alle c_k verschieden zu sein brauchen. Das zeigt das folgende Lemma.

Lemma. *Ist $A = \bigcup_k B_k$, $B_i \cap B_j = \emptyset$ für $i \neq j$ und nimmt die Funktion f auf B_k nur den Funktionswert c_k an, dann gilt*

$$\int_A f(x)\, d\mu = \sum_k c_k \mu(B_k), \tag{2}$$

und die Funktion f ist auf A genau dann integrierbar, wenn die Reihe (2) *absolut konvergiert.*

Beweis. Es ist leicht zu sehen, daß jede Menge

$$A_n = \{x : x \in A, f(x) = y_n\}$$

die Vereinigung aller Mengen B_k mit $c_k = y_n$ ist. Daraus folgt

$$\sum_n y_n \mu(A_n) = \sum_n y_n \Big(\sum_{c_k = y_n} \mu(B_k)\Big) = \sum_k c_k \mu(B_k).$$

Weil das Maß nichtnegativ ist, gilt

$$\sum_n |y_n|\, \mu(A_n) = \sum_n |y_n| \sum_{c_k = y_n} \mu(B_k) = \sum_k |c_k|\, \mu(B_k),$$

d. h., die Reihen $\sum_n y_n \mu(A_n)$ und $\sum_k c_k \mu(B_k)$ konvergieren absolut oder divergieren gleichzeitig. Damit ist das Lemma bewiesen.

Im weiteren stellen wir einige Eigenschaften des Lebesgueschen Integrals für Treppenfunktionen zusammen:

(A) *Es ist*

$$\int_A [f(x) + g(x)]\, d\mu = \int_A f(x)\, d\mu + \int_A g(x)\, d\mu,$$

wobei aus der Existenz der Integrale auf der rechten Seite die Existenz des Integrals auf der linken Seite folgt.

Zum Beweis setzen wir voraus, daß f die Werte f_i auf den Mengen $F_i \subset A$ und g die Werte g_j auf den Mengen $G_j \subset A$ annimmt, so daß

$$I_1 = \int_A f(x)\, d\mu = \sum_i f_i \mu(F_i), \tag{3}$$

$$I_2 = \int_A g(x)\, d\mu = \sum_j g_j \mu(G_j) \tag{4}$$

ist. Nach dem Lemma gilt

$$I = \int_A [f(x) + g(x)]\, d\mu = \sum_i \sum_j (f_i + g_j)\, \mu(F_i \cap G_j), \tag{5}$$

wobei sich aus

$$\mu(F_i) = \sum_j \mu(F_i \cap G_j), \qquad \mu(G_j) = \sum_i \mu(F_i \cap G_j)$$

und der absoluten Konvergenz der Reihen (3) und (4) die absolute Konvergenz der Reihe (5) und daraus

$$I = I_1 + I_2$$

ergibt.

(B) *Für eine beliebige Konstante k ist*

$$\int_A k f(x)\, d\mu = k \int_A f(x)\, d\mu,$$

wobei aus der Existenz des Integrals auf der rechten Seite die Existenz des Integrals auf der linken folgt.

Dieser Sachverhalt ergibt sich unmittelbar aus der Definition des Integrals.

(C) *Eine auf der Menge A beschränkte Treppenfunktion ist auf A integrierbar, wobei aus $|f(x)| \leqq M$, $x \in A$,*

$$\left| \int_A f(x)\, d\mu \right| \leqq M \mu(A)$$

folgt.

Dieser Sachverhalt ergibt sich ebenfalls unmittelbar aus der Definition des Integrals.

5.5.3. Die allgemeine Definition des Lebesgueschen Integrals auf einer Menge endlichen Maßes

Definition 3. Eine Funktion f heißt *integrierbar* (*summierbar*) auf der Menge A, wenn es eine Folge $\{f_n\}$ von integrierbaren Treppenfunktionen gibt, die gleichmäßig auf A gegen f konvergiert. Den Grenzwert

$$I = \lim_{n\to\infty} \int_A f_n(x)\, d\mu \tag{6}$$

bezeichnen wir mit

$$\int_A f(x)\, d\mu$$

und nennen ihn *Integral der Funktion f über die Menge A.*

Diese Definition ist korrekt, wenn folgende Bedingungen erfüllt sind:

1. Der Grenzwert (6) existiert für eine beliebige, auf A gleichmäßig konvergente Folge integrierbarer Treppenfunktionen.

2. Dieser Grenzwert hängt bei vorgegebenem f nicht von der Auswahl der Folge $\{f_n\}$ ab.

3. Für Treppenfunktionen stimmt diese Definition der Integrierbarkeit und des Integrals mit der in 5.5.2. angegebenen überein.

Alle diese Bedingungen sind tatsächlich erfüllt. Zum Beweis der ersten genügt es zu bemerken, daß auf Grund der Eigenschaften (A), (B), (C) des Integrals von Treppenfunktionen

$$\left|\int_A f_n(x)\, d\mu - \int_A f_m(x)\, d\mu\right| \leqq \mu(A) \cdot \sup_{x\in A} |f_n(x) - f_m(x)| \tag{7}$$

gilt. Für den Beweis der zweiten Bedingung betrachten wir zwei gegen f konvergente Folgen $\{f_n\}$ und $\{f_n^*\}$ mit den in der Definition genannten Eigenschaften. Nehmen wir an, daß der Grenzwert (6) dieser Folgen verschieden ist, dann existiert für die Folge $\{f_1, f_1^*, f_2, f_2^*, \ldots, f_n, f_n^*, \ldots\}$ kein Grenzwert (6). Das widerspricht der ersten Bedingung. Der Beweis für dritte Bedingung ergibt sich, wenn für die Treppenfunktion f die Folge $f_n(x) = f(x)$ $(n = 1, 2, \ldots)$ betrachtet wird.

Bemerkung. Die Konstruktion des Lebesgueschen Integrals geschieht also in zwei Etappen. 1. Etappe: Unmittelbare Definition des Integrals (als Summe einer Reihe) für eine hinreichend einfache und doch genügend umfangreiche Klasse von Funktionen (summierbare Treppenfunktionen). 2. Etappe: Ausdehnung der Definition des Integrals auf eine wesentlich umfangreichere Funktionenklasse mit Hilfe eines Grenzübergangs. Dieses Grundprinzip — Verknüpfung einer unmittelbar konstruktiven, aber engen Definition mit einem nachfolgenden Grenzübergang — findet sich im Grunde bei jeder beliebigen Konstruktion eines Integrals wieder.

Im folgenden stellen wir grundlegende Eigenschaften des Lebesgueschen Integrals zusammen.

I. *Es gilt*

$$\int_A 1 \cdot d\mu = \mu(A). \tag{8}$$

Das folgt unmittelbar aus der Definition.

II. *Für eine beliebige Konstante k ist*

$$\int_A k\, f(x)\, d\mu = k \int_A f(x)\, d\mu, \tag{9}$$

wobei aus der Existenz des Integrals auf der rechten Seite die Existenz des Integrals auf der linken folgt.

Diese Eigenschaft ergibt sich durch Grenzübergang aus der Eigenschaft (B) des Integrals von Treppenfunktionen.

III. *Additivität:*

$$\int_A [f(x) + g(x)]\, d\mu = \int_A f(x)\, d\mu + \int_A g(x)\, d\mu, \tag{10}$$

wobei aus der Existenz der Integrale auf der rechten Seite die Existenz des Integrals auf der linken folgt.

Der Beweis ergibt sich durch Grenzübergang aus der Eigenschaft (A) des Integrals von Treppenfunktionen.

IV. *Eine auf der Menge A beschränkte Funktion f ist auf A integrierbar.*

Der Beweis ergibt sich durch Grenzübergang aus der Eigenschaft (C) des Integrals von Treppenfunktionen unter Benutzung des Satzes 2.

V. *Monotonie: Wenn $f(x) \geqq 0$ ist, dann ist auch*

$$\int_A f(x)\, d\mu \geqq 0 \tag{11}$$

(unter der Voraussetzung, das Integral existiert).

Für Treppenfunktionen folgt diese Eigenschaft direkt aus der Definition. Im allgemeinen Fall ergibt sie sich aus der Tatsache, daß zu einer nichtnegativen meßbaren Funktion f auch eine gegen f gleichmäßig konvergente Folge nichtnegativer Treppenfunktionen konstruiert werden kann (vgl. Satz 2).

Aus der Eigenschaft V folgt sofort: Wenn $f(x) \geqq g(x)$ für alle (oder fast alle) $x \in A$ ist, dann ist auch

$$\int_A f(x)\, d\mu \leqq \int_A g(x)\, d\mu. \tag{12}$$

Damit gilt für $m \leqq f(x) \leqq M$

$$m\mu(A) \leqq \int_A f(x)\, d\mu \leqq M\mu(A) \tag{13}$$

VI. *Wenn* $\mu(A) = 0$ *ist, so ist* $\int\limits_A f(x)\,d\mu = 0$.

VI'. *Wenn* $f(x) = g(x)$ *fast überall auf* A *ist, dann ist auch*

$$\int\limits_A f(x)\,d\mu = \int\limits_A g(x)\,d\mu,$$

wobei beide Integrale gleichzeitig existieren oder nicht existieren.

Diese beiden Eigenschaften folgen unmittelbar aus der Definition des Lebesgueschen Integrals.

VII. *Wenn die Funktion* φ *auf* A *integrierbar und* $|f(x)| \leqq \varphi(x)$ *fast überall auf* A *ist, dann ist auch* f *auf* A *integrierbar.*

Sind nämlich f und φ Treppenfunktionen, dann läßt sich die Menge A in eine Menge vom Maß 0 und eine Restmenge A' zerlegen, so daß $A' = \bigcup\limits_n A_n'$ ist, die Funktionen f und φ auf den Mengen A_n' konstant sind ($f(x) = a_n$, $\varphi(x) = b_n$) und $|a_n| \leqq b_n$ ist. Aus der Integrierbarkeit von φ folgt $\sum\limits_n |a_n|\,\mu(A_n') \leqq \sum\limits_n b_n\,\mu(A_n')$ $= \int\limits_{A'} \varphi(x)\,d\mu = \int\limits_A \varphi(x)\,d\mu$, d. h., auch f ist integrierbar, und es gilt

$$\left|\int\limits_A f(x)\,d\mu\right| = \left|\int\limits_{A'} f(x)\,d\mu\right| = \left|\sum_n a_n\,\mu(A_n')\right|$$
$$\leqq \sum_n |a_n|\,\mu(A_n') = \int\limits_{A'} |f(x)|\,d\mu \leqq \int\limits_A \varphi(x)\,d\mu.$$

Mit Hilfe von Satz 2 ergibt sich aus dieser Aussage für Treppenfunktionen durch Grenzübergang die entsprechende Aussage für meßbare Funktion f und φ.

VIII. *Das Integral*

$$I_1 = \int\limits_A f(x)\,d\mu$$

existiert genau dann, wenn

$$I_2 = \int\limits_A |f(x)|\,d\mu \tag{14}$$

existiert.

Aus der Existenz des Integrals I_2 folgt nämlich nach Eigenschaft VII die Existenz von I_1. Die umgekehrte Folgerung ergibt sich für Treppenfunktionen unmittelbar aus der Definition des Integrals und für den allgemeinen Fall dann daraus mit Hilfe von Satz 2 und der Ungleichung

$$|a| - |b| \leqq |a - b|$$

durch Grenzübergang.

5.5.4. σ-Additivität und absolute Stetigkeit des Lebesgueschen Integrals. In 5.5.3. haben wir Eigenschaften des Lebesgueschen Integrals bei fester Menge dargestellt. Jetzt betrachten wir das Lebesguesche Integral als Mengenfunktion

$$F(A) = \int_A f(x)\,d\mu$$

auf der Gesamtheit aller meßbaren Mengen. Eine der wichtigsten Eigenschaften dieser Funktion ist folgende:

Satz 3. *Ist $A = \bigcup_n A_n$, $A_i \cap A_j = \emptyset$ für $i \neq j$, dann gilt*

$$\int_A f(x)\,d\mu = \sum_n \int_{A_n} f(x)\,d\mu, \tag{15}$$

wobei aus der Existenz des Integrals auf der linken Seite die Existenz der Integrale auf der rechten und die absolute Konvergenz der Reihe folgt.

Beweis. Zuerst beweisen wir die Behauptung des Satzes für Treppenfunktionen. Es sei f eine Treppenfunktion mit den Werten $y_1, y_2, \ldots, y_n, \ldots$, und wir setzen

$$B_k = \{x\colon x \in A,\ \ f(x) = y_k\},$$

$$B_{nk} = \{x\colon x \in A_n,\ f(x) = y_k\}.$$

Dann ist

$$\int_A f(x)\,d\mu = \sum_k y_k\,\mu(B_k) = \sum_k y_k \sum_n \mu(B_{nk})$$

$$= \sum_n \sum_k y_k\,\mu(B_{nk}) = \sum_n \int_{A_n} f(x)\,d\mu. \tag{16}$$

Weil die Reihe $\sum y_k\,\mu(B_k)$ auf Grund der vorausgesetzten Integrierbarkeit von f auf A absolut konvergiert und die Maße aller Mengen nichtnegativ sind, konvergieren auch alle restlichen Reihen in der Gleichungskette (16) absolut.

Im Fall einer beliebigen Funktion f folgt aus ihrer Integrierbarkeit auf A, daß es für jedes $\varepsilon > 0$ eine integrierbare Treppenfunktion g mit

$$|f(x) - g(x)| < \varepsilon, \quad x \in A, \tag{17}$$

gibt. Für g ist

$$\int_A g(x)\,d\mu = \sum_n \int_{A_n} g(x)\,d\mu, \tag{18}$$

wobei g nach den vorhergehenden Betrachtungen auf jeder Menge A_n integrierbar und die Reihe (18) absolut konvergent ist. Aus dieser Tatsache und der Abschätzung (17) folgt, daß auch f auf jedem A_n integrierbar und

$$\sum_n \left| \int_{A_n} f(x)\,d\mu - \int_{A_n} g(x)\,d\mu \right| \leqq \sum_n \varepsilon\mu(A_n) = \varepsilon\mu(A),$$

$$\left| \int_A f(x)\,d\mu - \int_A g(x)\,d\mu \right| \leqq \varepsilon\mu(A)$$

ist. Zusammen mit (18) ergeben sich daraus die absolute Konvergenz der Reihe $\sum_n \int_{A_n} f(x)\,d\mu$ und die Abschätzung

$$\left|\sum_n \int_{A_n} f(x)\,d\mu - \int_A f(x)\,d\mu\right| \leqq 2\varepsilon\mu(A).$$

Da $\varepsilon > 0$ beliebig gewählt werden kann, finden wir schließlich

$$\sum_n \int_{A_n} f(x)\,d\mu = \int_A f(x)\,d\mu.$$

Folgerung. *Wenn f auf A integrierbar ist, dann ist f auch auf einer beliebigen meßbaren Menge $A' \subset A$ integrierbar.*

Wir haben gezeigt, daß aus der Integrierbarkeit der Funktion f auf der Menge A für $A = \bigcup_n A_n$, $A_i \cap A_j = \emptyset$ $(i \neq j)$ die Integrierbarkeit von A auf jeder Menge A_n und die Darstellung des Integrals über A als Summe von Integralen über die Mengen A_n folgt. Diese Aussage kann in folgendem Sinn umgekehrt werden.

Satz 4. *Ist $A = \bigcup_n A_n$, $A_i \cap A_j = \emptyset$ für $i \neq j$, und konvergiert die Reihe*

$$\sum_n \int_{A_n} |f(x)|\,d\mu, \tag{19}$$

dann ist die Funktion f integrierbar auf A und

$$\int_A f(x)\,d\mu = \sum_n \int_{A_n} f(x)\,d\mu.$$

Beweis. Neu ist hier im Vergleich zum vorhergehenden Satz lediglich die Behauptung, daß aus der Konvergenz der Reihe (19) die Integrierbarkeit von f auf A folgt.

Wir führen den Beweis zunächst wieder für den Fall einer Treppenfunktion f mit den Werten f_i. Setzen wir

$$B_i = \{x : x \in A, f(x) = f_i\}, \qquad A_{ni} = A_n \cap B_i,$$

so ist

$$\bigcup_n A_{ni} = B_i \quad \text{und} \quad \int_{A_n} |f(x)|\,d\mu = \sum_i |f_i|\,\mu(A_{ni}).$$

Aus der Konvergenz der Reihe (19) folgt nun die Konvergenz der Reihen

$$\sum_n \sum_i |f_i|\,\mu(A_{ni}) = \sum_i |f_i|\,\mu(B_i),$$

d. h., es existiert das Integral

$$\int_A f(x)\,d\mu = \sum_i f_i\,\mu(B_i).$$

Im allgemeinen Fall approximieren wir f durch eine Treppenfunktion $\tilde{f}$:

$$|f(x) - \tilde{f}(x)| < \varepsilon, \qquad x \in A. \tag{20}$$

Damit ergibt sich

$$\int_{A_n} |\tilde{f}(x)| \, d\mu \leqq \int_{A_n} |f(x)| \, d\mu + \varepsilon\mu(A_n),$$

und weil die Reihe

$$\sum_n \mu(A_n) = \mu(A)$$

konvergiert, folgt aus der Konvergenz der Reihe (19) die Konvergenz der Reihe

$$\sum_n \int_{A_n} |\tilde{f}(x)| \, d\mu.$$

Nach der für Treppenfunktionen bereits bewiesenen Aussage des Satzes ist $\tilde{f}$ auf A integrierbar. Auf Grund von (20) ist damit auch f auf A integrierbar, was zu zeigen war.

Čebyševsche Ungleichung. Wenn $\varphi(x) \geqq 0$ und integrierbar auf A ist, dann gilt für eine beliebige positive Zahl c die Ungleichung

$$\mu\{x\colon x \in A,\, \varphi(x) \geqq c\} \leqq \frac{1}{c} \int_A \varphi(x) \, d\mu. \tag{21}$$

Zum Beweis setzen wir

$$A' = \{x\colon x \in A,\, \varphi(x) \geqq c\}.$$

Es ergibt sich

$$\int_A \varphi(x) \, d\mu = \int_{A'} \varphi(x) \, d\mu + \int_{A \setminus A'} \varphi(x) \, d\mu \geqq \int_{A'} \varphi(x) \, d\mu \geqq c\mu(A').$$

Folgerung. *Ist*

$$\int_A |f(x)| \, d\mu = 0,$$

dann ist $f(x) = 0$ fast überall auf A.

Aus der Voraussetzung folgt mit der Čebyševschen Ungleichung

$$\mu\left\{x\colon x \in A,\, |f(x)| \geqq \frac{1}{n}\right\} \leqq n \int_A |f(x)| \, d\mu = 0$$

für jedes $n > 0$ und daher aus

$$\mu\{x\colon x \in A,\, f(x) \neq 0\} \leqq \sum_{n=1}^{\infty} \mu\left\{x\colon x \in A,\, |f(x)| \geqq \frac{1}{n}\right\} = 0$$

die Behauptung.

In 5.5.3. wurde gezeigt, daß das Lebesguesche Integral über eine Menge vom Maß Null für eine beliebige Funktion f gleich Null ist. Diese Aussage kann man als Grenzfall des folgenden wichtigen Satzes auffassen.

Satz 5 (Absolute Stetigkeit des Lebesgueschen Integrals). *Wenn die Funktion $f(x)$ auf der Menge A integrierbar ist, dann existiert für jedes $\varepsilon > 0$ ein $\delta > 0$, so daß*

$$\left| \int_e f(x)\, d\mu \right| < \varepsilon$$

für jede meßbare Menge $e \subset A$ gilt, deren Maß $\mu(e) < \delta$ ist.

Beweis. Für eine beschränkte Funktion f ist die Behauptung offensichtlich. Ist f eine beliebige auf A integrierbare Funktion, so setzen wir

$$A_n = \{x : x \in A,\ n < |f(x)| \leqq n + 1\}$$

und

$$B_N = \bigcup_{n=0}^{N} A_n, \qquad C_N = A \setminus B_N.$$

Aus Satz 3 folgt

$$\int_A |f(x)|\, d\mu = \sum_{n=0}^{\infty} \int_{A_n} |f(x)|\, d\mu.$$

Wählen wir nun N so, daß

$$\sum_{n=N+1}^{\infty} \int_{A_n} |f(x)|\, d\mu = \int_{C_N} |f(x)|\, d\mu < \frac{\varepsilon}{2}$$

ist, und eine Zahl δ mit der Eigenschaft

$$0 < \delta < \frac{\varepsilon}{2(N+1)};$$

dann erfüllt diese Zahl δ die Forderungen des Satzes. Denn für eine beliebige Menge $e \subset A$ mit $\mu(e) < \delta$ folgt aus

$$\left| \int_e f(x)\, d\mu \right| \leqq \int_e |f(x)|\, d\mu = \int_{e \cap B_N} |f(x)|\, d\mu + \int_{e \cap C_N} |f(x)|\, d\mu$$

und aus

$$\int_{e \cap B_N} |f(x)|\, d\mu < \frac{\varepsilon}{2} \qquad \text{(nach Eigenschaft V)},$$

$$\int_{e \cap C_N} |f(x)|\, d\mu \leqq \int_{C_N} |f(x)|\, d\mu < \frac{\varepsilon}{2}$$

die nachzuprüfende Eigenschaft

$$\left|\int_e f(x)\,d\mu\right| < \varepsilon.$$

Fassen wir alle angegebenen Eigenschaften des Integrals als Mengenfunktion zusammen, dann ergibt sich folgendes Resultat: *Ist f eine nichtnegative, auf der ganzen Menge X bezüglich des Maßes μ integrierbare Funktion, dann ist die Funktion*

$$F(A) = \int_A f(x)\,d\mu$$

für alle meßbaren Mengen $A \subset X$ definiert, nichtnegativ und σ-additiv, d. h., es gilt

$$F(A) = \sum_n F(A_n),$$

wenn $A = \bigcup_n A_n$ und $A_i \cap A_j = \emptyset$ $(i \neq j)$ ist. Mit anderen Worten, *das Integral einer nichtnegativen Funktion besitzt als Mengenfunktion alle Eigenschaften eines σ-additiven Maßes.* Dieses Maß ist auf derselben σ-Algebra wie das Ausgangsmaß μ definiert und mit μ über die Bedingung „Aus $\mu(A) = 0$ folgt $F(A) = 0$“ verknüpft.

5.5.5. Grenzübergang unter dem Lebesgueschen Integralzeichen. Die Frage, ob ein Grenzübergang unter das Integralzeichen gezogen werden oder eine konvergente Reihe gliedweise integriert werden kann, d. h. die Frage nach der Vertauschbarkeit von Grenzwert- und Integralbildung tritt bei den verschiedensten Problemen auf. In der klassischen Analysis wird dafür als hinreichende Bedingung die gleichmäßige Konvergenz der entsprechenden Folge (Reihe) angegeben. Wir beweisen nun einige Sätze über den Grenzübergang unter dem Lebesgueschen Integralzeichen, die weitgehende Verallgemeinerungen der entsprechenden Sätze der klassischen Analysis sind.

Satz 6 (Lebesgue). *Konvergiert die Folge $\{f_n\}$ auf A gegen f und existiert eine auf A integrierbare Funktion $\varphi(x)$, so daß*

$$|f_n(x)| \leqq \varphi(x), \quad x \in A,$$

für alle n gilt, dann ist die Grenzfunktion f auf A integrierbar, und es gilt

$$\int_A f_n(x)\,d\mu \to \int_A f(x)\,d\mu.$$

Beweis. Aus den Voraussetzungen des Satzes ergibt sich sofort, daß $|f(x)| \leqq \varphi(x)$ ist. Nach Eigenschaft VIII aus 5.5.3. ist dann f integrierbar. Setzen wir

$$A_k = \{x\colon k \leqq \varphi(x) < k+1\}, \qquad k = 0, 1, 2, \ldots,$$

$$B_m = \bigcup_{k \geqq m} A_k = \{x\colon \varphi(x) \geqq m\}, \qquad m = 0, 1, 2, \ldots,$$

so folgt aus Satz 3

$$\int_A \varphi(x)\,d\mu = \sum_k \int_{A_k} \varphi(x)\,d\mu, \tag{22}$$

wobei die Reihe (22) absolut konvergiert. Wegen

$$\int_{B_m} \varphi(x)\,d\mu = \sum_{k \geqq m} \int_{A_k} \varphi(x)\,d\mu$$

existiert auf Grund der Konvergenz der Reihe (22) eine Zahl m, so daß

$$\int_{B_m} \varphi(x)\,d\mu < \frac{\varepsilon}{5}$$

ist. Auf $A \setminus B_m$ gilt $\varphi(x) < m$. Nach dem Satz von Egorov ist die Menge $A \setminus B_m$ in einer Form $A \setminus B_m = C \cup D$ darstellbar, wobei $\mu(D) < \varepsilon/5m$ ist, und die Folge $\{f_n\}$ auf C gleichmäßig gegen f konvergiert. Wählen wir nun N so groß, daß für $n > N$ auf C

$$|f_n(x) - f(x)| < \frac{\varepsilon}{5\mu(C)}$$

ist, so folgt aus der Zerlegung

$$\int_A [f_n(x) - f(x)]\,d\mu = \int_{B_m} f_n(x)\,d\mu - \int_{B_m} f(x)\,d\mu + \int_D f_n(x)\,d\mu - \int_D f(x)\,d\mu + \int_C [f_n(x) - f(x)]\,d\mu$$

die Abschätzung

$$\int_A |f_n(x) - f(x)|\,d\mu < 5 \cdot \frac{\varepsilon}{5} = \varepsilon.$$

Damit ist der Satz bewiesen.

Folgerung. *Im Fall* $|f_n(x)| \leqq M = \text{const}$ *und* $f_n \to f$ *gilt*

$$\int_A f_n(x)\,d\mu \to \int_A f(x)\,d\mu.$$

Bemerkung. Da Funktionswerte, die auf einer Menge von Maß Null angenommen werden, keinen Einfluß auf Größe des Integrals haben, genügt es, in Satz 6 vorauszusetzen, daß $\{f_n\}$ fast überall gegen f konvergiert und jede der Ungleichungen $|f_n(x)| \leqq \varphi(x)$ nur fast überall erfüllt ist.

Satz 7 (B. Levi). *Ist auf A eine nichtfallende Folge*

$$f_1(x) \leqq f_2(x) \leqq \cdots \leqq f_n(x) \leqq \cdots$$

integrierbarer Funktionen gegeben, deren Integrale gleichmäßig beschränkt sind,

$$\int_A f_n(x)\,d\mu \leqq K,$$

dann existiert fast überall auf A ein (endlicher) Grenzwert

$$f(x) = \lim_{n\to\infty} f_n(x). \tag{23}$$

Die Funktion f ist integrierbar auf A, und es gilt

$$\int_A f_n(x)\, d\mu \to \int_A f(x)\, d\mu.$$

Dabei kann die Funktion $f(x)$ für die x, in denen der Grenzwert (23) nicht existiert, beliebig festgesetzt werden, z. B. $f(x) = 0$.

Beweis. Wir können $f_1(x) \geqq 0$ voraussetzen, da der allgemeine Fall auf diesen mit Hilfe der Substitution

$$\bar{f}_n = f_n - f_1$$

zurückgeführt werden kann. Ist

$$\Omega = \{x : x \in A,\ f_n(x) \to \infty\},$$

so ergibt sich leicht

$$\Omega = \bigcap_r \bigcup_n \Omega_n^{(r)}, \qquad \Omega_n^{(r)} = \{x : x \in A,\ f_n(x) > r\}.$$

Nach der Čebyševschen Ungleichung (21) gilt

$$\mu(\Omega_n^{(r)}) \leqq \frac{K}{r}$$

und wegen $\Omega_1^{(r)} \subset \Omega_2^{(r)} \subset \cdots \subset \Omega_n^{(r)} \subset \cdots$ auch

$$\mu\Big(\bigcup_n \Omega_n^{(r)}\Big) \leqq \frac{K}{r}.$$

Da für beliebiges r

$$\Omega \subset \bigcup_n \Omega_n^{(r)}$$

ist, folgt $\mu(\Omega) \leqq \frac{K}{r}$, d. h. für frei wählbares r

$$\mu(\Omega) = 0.$$

Damit ist gezeigt, daß die monoton wachsende Folge $\{f_n\}$ fast überall auf A einen endlichen Grenzwert besitzt.

Bezeichnen wir mit A_r die Menge

$$A_r = \{x : x \in A,\ r - 1 \leqq f(x) < r\}, \qquad r = 1, 2, \ldots,$$

und setzen $\varphi(x) = r$ auf A, so genügt es für den Beweis des Satzes, die Integrierbarkeit von $\varphi(x)$ auf A_r nachzuweisen. Denn dann folgt die Behauptung unmittelbar aus

Satz 6. Ist

$$B_s = \bigcup_{r=1}^{s} A_r,$$

so sind die Funktionen f_n und f auf B_s beschränkt, und es gilt $\varphi(x) \leqq f(x) + 1$. Daraus ergibt sich

$$\int_{B_s} \varphi(x)\,d\mu \leqq \int_{B_s} f(x)\,d\mu + \mu(A) = \lim_{n\to\infty} \int_{B_s} f_n(x)\,d\mu + \mu(A) \leqq K + \mu(A).$$

Andererseits ist

$$\int_{B_s} \varphi(x)\,d\mu = \sum_{r=1}^{s} r\mu(A_r).$$

Die Beschränktheit aller dieser Summen liefert die Konvergenz der Reihe

$$\sum_{r=1}^{\infty} r\mu(A_r) = \int_A \varphi(x)\,d\mu,$$

d. h. die Integrierbarkeit der Funktion φ auf A, was zu zeigen war.

Die Voraussetzung, daß die Folge $\{f_n\}$ nichtfallend ist, kann offensichtlich durch die Bedingung, daß $\{f_n\}$ nichtwachsend ist, ersetzt werden. Es gilt dann ein entsprechender Satz.

Folgerung. *Wenn $\psi_n(x) \geqq 0$ für alle n und*

$$\sum_{n=1}^{\infty} \int_A \psi_n(x)\,d\mu < \infty$$

ist, dann konvergiert die Reihe $\sum_{n=1}^{\infty} \psi_n(x)$ fast überall auf A, und es gilt

$$\int_A \left(\sum_{n=1}^{\infty} \psi_n(x)\right) d\mu = \sum_{n=1}^{\infty} \int_A \psi_n(x)\,d\mu.$$

Satz 8 (Fatou). *Konvergiert eine Folge $\{f_n\}$ meßbarer nichtnegativer Funktionen fast überall auf A gegen die Funktion f und ist für alle n*

$$\int_A f_n(x)\,d\mu \leqq K,$$

dann ist f auf A integrierbar und

$$\int_A f(x)\,d\mu \leqq K.$$

Beweis. Setzen wir

$$\varphi_n(x) = \inf_{k \geqq n} f_k(x),$$

so ist φ_n meßbar wegen

$$\{x: \varphi_n(x) < c\} = \bigcup_{k \geqq n} \{x: f_k(x) < c\}.$$

Aus $0 \leqq \varphi_n(x) \leqq f_n(x)$ folgt die Integrierbarkeit der φ_n und

$$\int_A \varphi_n(x)\, d\mu \leqq \int_A f_n(x)\, d\mu \leqq K.$$

Diese Beziehung und die offensichtlich für die φ_n bestehenden Eigenschaften

$$\varphi_1(x) \leqq \varphi_2(x) \leqq \cdots \leqq \varphi_n(x) \leqq \cdots,$$

$$\lim_{n\to\infty} \varphi_n(x) = f(x) \text{ fast überall}$$

zeigen, daß für die Folge $\{\varphi_n\}$ die Voraussetzungen von Satz 7 erfüllt sind. Wenden wir Satz 7 auf die Folge $\{\varphi_n\}$ an, so ergibt sich die Aussage von Satz 8.

5.5.6. Das Lebesguesche Integral über eine Menge unendlichen Maßes. Wenn wir bisher über das Integral und seine Eigenschaften sprachen, nahmen wir stets an, daß die betrachteten Funktionen auf irgendeiner meßbaren Menge endlichen Maßes gegeben waren. Es erweist sich jedoch oft als notwendig, Funktionen zu betrachten, die auf einer Menge unendlichen Maßes gegeben sind, z. B. auf der Geraden mit dem Lebesgueschen Maß. Daher ist es wichtig, den Begriff des Integrals auch auf diesen Fall auszudehnen. Wir beschränken uns dabei auf den praktisch wichtigsten Fall, in dem die betrachtete Menge X als abzählbare Vereinigung von Mengen endlichen Maßes

$$X = \bigcup_n X_n, \quad \mu(X_n) < \infty, \tag{24}$$

dargestellt werden kann.

Wenn die Menge X, auf der das Maß μ gegeben ist, als abzählbare Vereinigung von Mengen endlichen Maßes darstellbar ist, dann heißt das Maß μ σ-endlich auf X (siehe 5.3.3.). Beispiele für σ-endliche Maße sind das Lebesguesche Maß auf der Geraden, in der Ebene und im n-dimensionalen Raum. Ein Maß, das nicht die Bedingung der σ-Endlichkeit erfüllt, ist z. B. die Funktion, die jedem Punkt der Geraden das Maß Eins zuordnet. Dabei können wir alle Untermengen der Geraden als meßbar annehmen, wobei die endlichen Mengen ein entsprechendes endliches Maß haben und die restlichen Mengen von unendlichem Maß sind.

Jede Folge $\{X_n\}$ von meßbaren Untermengen der Menge X, die der Bedingung (24) genügen, bezeichnen wir als *ausschöpfende Folge*. Unter Verwendung solcher Folgen definieren wir den allgemeinen Integralbegriff.

Definition 4. Auf der Menge X sei ein σ-endliches Maß μ gegeben. Ist f eine auf X definierte meßbare Funktion, so heißt f *integrierbar* (*summierbar*) auf X, wenn f auf jeder meßbaren Untermenge $A \subset X$ integrierbar (summierbar) ist und für jede

ausschöpfende Folge $\{X_n\}$ der Grenzwert

$$\lim_{n\to\infty} \int_{X_n} f(x)\, d\mu \tag{25}$$

existiert. Dieser Grenzwert heißt dann *Integral von f über die Menge X* und wird mit dem Symbol

$$\int_X f(x)\, d\mu$$

bezeichnet.

Der Grenzwert (25) hängt offenbar nicht von der Auswahl der ausschöpfenden Folge $\{X_n\}$ ab. Denn nehmen wir an, daß dieser Grenzwert für zwei ausschöpfende Folgen $\{X_n'\}$ und $\{X_n''\}$ verschieden wäre, so würde er für die Folge $\{X_1', X_1'', X_2', X_2'', \ldots\}$, die ebenfalls ausschöpfend ist, im Widerspruch zur Voraussetzung nicht existieren. Offensichtlich stimmt auch die obige Definition des Integrals mit der in 5.5.3. angegebenen überein, wenn die Funktion f außerhalb einer Menge endlichen Maßes gleich 0 ist.

Bemerkung. Die Definition des Integrals für Treppenfunktionen, wie sie in 5.5.2. angegeben wurde, kann wörtlich auf den Fall von Mengen unendlichen Maßes übertragen werden. Dafür, daß eine Treppenfunktion summierbar ist, ist es dann offensichtlich notwendig, daß von Null verschiedene Funktionswerte nur auf Mengen endlichen Maßes angenommen werden. Die Definition der Integrierbarkeit, wie sie in 5.5.3. angegeben wurde, ist nur unter der Voraussetzung der Endlichkeit des Maßes der Menge X sinnvoll. Denn für X mit $\mu(X) = \infty$ folgt aus der gleichmäßigen Konvergenz einer Folge von summierbaren Treppenfunktionen im allgemeinen nicht die Konvergenz der Folge ihrer Integrale. (Der Leser gebe ein Beispiel an!)

Die Resultate, die in 5.5.3. und 5.5.4. für den Fall eines endlichen Maßes angegeben wurden, lassen sich im wesentlichen auf Integrale über Mengen unendlichen Maßes übertragen.

Der Unterschied zwischen beiden Integraldefinitionen besteht vor allem darin, daß im Fall $\mu(X) = \infty$ eine beschränkte meßbare Funktion nicht mehr notwendig integrierbar sein muß. Insbesondere ist für $\mu(X) = \infty$ keine einzige von 0 verschiedene Konstante auf X integrierbar.

Der Leser prüft ohne Mühe nach, daß die Sätze von Lebesgue, B. Levi und Fatou auch im Fall von Mengen unendlichen Maßes richtig sind.

5.5.7. Vergleich des Lebesgueschen Integrals mit dem Riemannschen Integral. Wir klären nun den Zusammenhang zwischen den Integralen von Lebesgue und Riemann auf. Dabei beschränken wir uns auf den einfachsten Fall des Lebesgueschen Maßes auf der Geraden.

Satz 9. *Wenn für eine Funktion f das Riemannsche Integral*

$$I = (R)\int_a^b f(x)\,dx$$

existiert, dann ist f auf $[a, b]$ integrierbar (nach Lebesgue), und es gilt

$$\int_{[a,b]} f(x)\,d\mu = I.$$

Beweis. Wir zerlegen das Intervall $[a, b]$ durch die Punkte

$$x_k = a + \frac{k}{2^n}(b - a)$$

in 2^n Teile und betrachten die dieser Zerlegung entsprechenden Darbouxschen Summen:

$$\Omega_n = \frac{b-a}{2^n}\sum_{k=1}^{2^n} M_{nk},$$

$$\omega_n = \frac{b-a}{2^n}\sum_{k=1}^{2^n} m_{nk}.$$

Dabei ist M_{nk} die obere Grenze und m_{nk} die untere Grenze der Funktionswerte $f(x)$ auf dem Intervall

$$x_{k-1} \leqq x \leqq x_k.$$

Nach Definition des Riemannschen Integrals ist

$$I = \lim_{n\to\infty} \Omega_n = \lim_{n\to\infty} \omega_n.$$

Setzen wir

$$\bar{f}_n(x) = M_{nk} \quad \text{für } x_{k-1} \leqq x < x_k,$$

$$\underline{f}_n(x) = m_{nk} \quad \text{für } x_{k-1} \leqq x < x_k$$

und legen im Punkt $x = b$ für die Funktionen $\bar{f}_n$ und $\underline{f}_n$ einen beliebigen Funktionswert fest, so ist leicht auszurechnen, daß

$$\int_{[a,b]} \bar{f}_n(x)\,d\mu = \Omega_n,$$

$$\int_{[a,b]} \underline{f}_n(x)\,d\mu = \omega_n$$

ist. Weil die Folge $\{\bar{f}_n\}$ nicht wächst und die Folge $\{\underline{f}_n\}$ nicht fällt, gilt fast überall auf $[a, b]$

$$\bar{f}_n(x) \to \bar{f}(x) \geqq f(x), \qquad \underline{f}_n(x) \to \underline{f}(x) \leqq f(x).$$

Aus dem Satz von B. Levi folgt

$$\int\limits_{[a,b]} \bar{f}(x)\, d\mu = \lim_{n\to\infty} \Omega_n = I = \lim_{n\to\infty} \omega_n = \int\limits_{[a,b]} \underline{f}(x)\, d\mu\,,$$

d. h.

$$\int\limits_{[a,b]} |\bar{f}(x) - \underline{f}(x)|\, d\mu = \int\limits_{[a,b]} \big(\bar{f}(x) - \underline{f}(x)\big)\, d\mu = 0.$$

Nach der Folgerung aus der Čebyševschen Ungleichung ist fast überall auf $[a, b]$

$$\bar{f}(x) - \underline{f}(x) = 0,$$

d. h.

$$\bar{f}(x) = \underline{f}(x) = f(x)$$

und damit

$$\int\limits_{[a,b]} f(x)\, d\mu = I,$$

was zu zeigen war.

Man kann leicht Beispiele beschränkter Funktionen angeben, die auf einem Intervall integrierbar nach Lebesgue, aber nicht integrierbar nach Riemann sind (z. B. die Dirichletsche Funktion, die für rationale x aus $[0, 1]$ gleich 1 und für irrationale x aus $[0, 1]$ gleich 0 ist). Unbeschränkte Funktionen sind nach Riemann (im eigentlichen Sinn) überhaupt nicht integrierbar, während viele von ihnen nach Lebesgue integrierbar sind. Insbesondere ist eine beliebige Funktion $f(x) \geqq 0$, deren Riemannsches Integral

$$\int\limits_{a+\varepsilon}^{b} f(x)\, dx$$

für jedes $\varepsilon > 0$ existiert und einen endlichen Grenzwert für $\varepsilon \to 0$ besitzt, auf $[a, b]$ nach Lebesgue integrierbar, und es gilt

$$\int\limits_{[a,b]} f(x)\, d\mu = \lim_{\varepsilon\to 0} \int\limits_{a+\varepsilon}^{b} f(x)\, dx.$$

Das uneigentliche Riemannsche Integral

$$\lim_{\varepsilon\to 0} \int\limits_{a+\varepsilon}^{b} f(x)\, dx$$

existiert nicht im Lebesgueschen Sinne, wenn

$$\lim_{a\to 0} \int\limits_{a+\varepsilon}^{b} |f(x)|\, dx = \infty$$

ist. Denn nach Eigenschaft VIII des Lebesgueschen Integrals (siehe 5.5.3.) folgt aus der Summierbarkeit von $f(x)$ die Summierbarkeit von $|f(x)|$. Beispielsweise existiert

$$\int_0^1 \frac{1}{x} \sin \frac{1}{x}\, dx$$

als (bedingt konvergentes) uneigentliches Riemannsches Integral. Es existiert aber nicht als Lebesguesches Integral.

Wird eine Funktion auf der ganzen Achse (oder Halbachse) betrachtet, dann kann das Riemannsche Integral für diese Funktion nur im uneigentlichen Sinn existieren. Ist dieses Integral absolut konvergent, so existiert (wie oben) das entsprechende Lebesguesche Integral und besitzt den gleichen Wert. Konvergiert aber dieses Integral nur bedingt, dann ist die Funktion im Lebesgueschen Sinn nicht integrierbar. So ist z. B. die Funktion

$$\frac{\sin x}{x}$$

auf der ganzen Achse nicht nach Lebesgue integrierbar, weil

$$\int_{-\infty}^{\infty} \left| \frac{\sin x}{x} \right| dx = \infty$$

ist. Das uneigentliche Riemannsche Integral

$$\int_{-\infty}^{\infty} \frac{\sin x}{x}\, dx$$

existiert jedoch und hat den Wert π.

5.6. Direkte Produkte von Mengensystemen und Maßen. Der Satz von Fubini.

In der Analysis spielen die Sätze über die Zurückführung eines zweifachen (oder allgemein mehrfachen) Integrals auf ein iteriertes Integral eine wichtige Rolle. In der Theorie der mehrfachen Lebesgueschen Integrale wird das grundlegende Resultat in dieser Hinsicht durch den sogenannten Satz von Fubini beschrieben, der am Ende dieses Abschnitts bewiesen wird. Vorbereitend stellen wir einige Begriffe und Fakten zusammen, die auch für sich von Interesse sind.

5.6.1. Produkte von Mengensystemen. Die Menge Z aller geordneten Paare (x, y), wobei $x \in X$ und $y \in Y$ ist, heißt *direktes Produkt* der Mengen X und Y und wird mit $X \times Y$ bezeichnet. Analog heißt die Menge aller geordneten n-Tupel $(x_1, x_2, \ldots, x_n)$, wobei $x_k \in X_k$ ist, direktes Produkt der Mengen $X_1, X_2, \ldots, X_n$ und wird mit

$$Z = X_1 \times X_2 \times \cdots \times X_n = \bigtimes_k X_k$$

bezeichnet. Wenn speziell

$$X_1 = X_2 = \cdots = X_n = X$$

gilt, ist Z die n-te Potenz der Menge X:

$$Z = X^n.$$

So ist z. B. der n-dimensionale Koordinatenraum $\boldsymbol{R}^n$ die n-te Potenz der Zahlengeraden $\boldsymbol{R}^1$. Der Einheitswürfel I^n, d. h. die Menge aller Elemente aus $\boldsymbol{R}^n$, deren Koordinaten die Bedingung

$$0 \leqq x_k \leqq 1, \quad k = 1, 2, \ldots, n,$$

erfüllen, ist die n-te Potenz des Einheitsintervalls $I^1 = [0, 1]$.

Wenn $\mathfrak{S}_1, \mathfrak{S}_2, \ldots, \mathfrak{S}_n$ Mengensysteme von Untermengen der Mengen $X_1, X_2, \ldots, X_n$ sind, bezeichnet

$$\mathfrak{R} = \mathfrak{S}_1 \times \mathfrak{S}_2 \times \cdots \times \mathfrak{S}_n = \bigtimes_k \mathfrak{S}_k$$

das System derjenigen Untermengen der Menge $X = \bigtimes_k X_k$, die in der Form

$$A = A_1 \times A_2 \times \cdots \times A_n$$

mit $A_k \in \mathfrak{S}_k$ darstellbar sind.

Im Fall $\mathfrak{S}_1 = \mathfrak{S}_2 = \cdots = \mathfrak{S}_n = \mathfrak{S}$ ist $\mathfrak{R}$ die n-te Potenz des Systems $\mathfrak{S}$:

$$\mathfrak{R} = \mathfrak{S}^n.$$

Zum Beispiel ist das System der Parallelepipede in $\boldsymbol{R}^n$ die n-te Potenz des Systems aller Intervalle in $\boldsymbol{R}^1$.

Satz 1. *Sind $\mathfrak{S}_1, \mathfrak{S}_2, \ldots, \mathfrak{S}_n$ Semiringe, dann ist auch $\mathfrak{R} = \bigtimes_k \mathfrak{S}_k$ ein Semiring.*

Beweis. Entsprechend der Definition des Semiringes müssen wir zeigen, daß für $A, B \in \mathfrak{R}$ stets $A \cap B \in \mathfrak{R}$ ist und für $B \subset A$ eine Darstellung von A in der Form $A = \bigcup_{i=1}^{m} C_i$ existiert, wobei $C_1 = B$, $C_i \cap C_j = \emptyset$ für $i \neq j$, $C_i \in \mathfrak{R}$ $(i = 1, 2, \ldots, n)$ ist. Wir führen den Beweis für den Fall $n = 2$.

a) Es sei $A \in \mathfrak{S}_1 \times \mathfrak{S}_2$, $B \in \mathfrak{S}_1 \times \mathfrak{S}_2$; d. h.

$$A = A_1 \times A_2, \quad A_1 \in \mathfrak{S}_1,\ A_2 \in \mathfrak{S}_2,$$
$$B = B_1 \times B_2, \quad B_1 \in \mathfrak{S}_1,\ B_2 \in \mathfrak{S}_2.$$

Dann ist

$$A \cap B = (A_1 \cap B_1) \times (A_2 \cap B_2),$$

und weil

$$A_1 \cap B_1 \in \mathfrak{S}_1, \quad A_2 \cap B_2 \in \mathfrak{S}_2$$

ist, folgt

$$A \cap B \in \mathfrak{S}_1 \times \mathfrak{S}_2.$$

b) Setzen wir zusätzlich voraus, daß $B \subset A$, d.h. $B_1 \subset A_1$, $B_2 \subset A_2$ ist, so gibt es für A_1 und A_2 in den Semiringen $\mathfrak{S}_1$ und $\mathfrak{S}_2$ Zerlegungen der Form

$$A_1 = B_1 \cup B_1^{(1)} \cup \cdots \cup B_1^{(k)},$$

$$A_2 = B_2 \cup B_2^{(1)} \cup \cdots \cup B_2^{(l)}.$$

Daraus ergibt sich für A eine Zerlegung

$$\begin{aligned} A = A_1 \times A_2 = (B_1 \times B_2) &\cup (B_1 \times B_2^{(1)}) \cup \cdots \cup (B_1 \times B_2^{(l)}) \\ &\cup (B_1^{(1)} \times B_2) \cup (B_1^{(1)} \times B_2^{(1)}) \cup \cdots \cup (B_1^{(1)} \times B_2^{(l)}) \\ &\quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &\cup (B_1^{(k)} \times B_2) \cup (B_1^{(k)} \times B_2^{(1)}) \cup \cdots \cup (B_1^{(k)} \times B_2^{(l)}), \end{aligned}$$

in der das erste Glied $B_1 \times B_2 = B$ ist und alle Glieder zu $\mathfrak{S}_1 \times \mathfrak{S}_2$ gehören. Damit ist der Satz bewiesen.

Aus der Voraussetzung, daß die Systeme $\mathfrak{S}_k$ Ringe (bzw. σ-Algebren) sind, folgt im allgemeinen nicht, daß das Produkt $\mathop{\times}\limits_k \mathfrak{S}_k$ ein Ring (bzw. eine σ-Algebra) ist.

5.6.2. Produktmaße. Auf den Semiringen $\mathfrak{S}_1, \mathfrak{S}_2, \ldots, \mathfrak{S}_n$ seien die Maße

$$\mu_1(A_1), \mu_2(A_2), \ldots, \mu_n(A_n), \quad A_k \in \mathfrak{S}_k,$$

gegeben. Dann definieren wir auf dem Semiring

$$\mathfrak{R} = \mathfrak{S}_1 \times \mathfrak{S}_2 \times \cdots \times \mathfrak{S}_n \tag{1}$$

das Maß

$$\mu = \mu_1 \times \mu_2 \times \cdots \times \mu_n \tag{2}$$

durch die Formel

$$\mu(A) = \mu_1(A_1)\, \mu_2(A_2) \cdots \mu_n(A_n). \tag{3}$$

Um zu zeigen, daß diese Festsetzung wirklich ein Maß (im Sinne 5.2.1.) definiert, genügt es, die Additivität dieser Mengenfunktion nachzuweisen. Wir führen diesen

Nachweis für den Fall $n = 2$. Es sei

$$A = A_1 \times A_2 = \bigcup_k B^{(k)}, \quad B^{(i)} \cap B^{(j)} = \emptyset \text{ für } i \neq j, \quad B^{(k)} = B_1^{(k)} \times B_2^{(k)}.$$

Nach Lemma 2 aus 1.5. existieren Zerlegungen

$$A_1 = \bigcup_m C_1^{(m)}, \quad A_2 = \bigcup_n C_2^{(n)},$$

so daß die Mengen $B_1^{(k)}$ als Vereinigung gewisser $C_1^{(m)}$ und die Mengen $B_2^{(k)}$ als Vereinigung gewisser $C_2^{(n)}$ dargestellt werden können. Dann ist offensichtlich

$$\mu(A) = \mu_1(A_1)\, \mu_2(A_2) = \sum_m \sum_n \mu_1(C_1^{(m)})\, \mu_2(C_2^{(n)}), \tag{4}$$

$$\mu(B^{(k)}) = \mu_1(B_1^{(k)})\, \mu_2(B_2^{(k)}) = \sum_m \sum_n \mu_1(C_1^{(m)})\, \mu_2(C_2^{(n)}), \tag{5}$$

wobei auf der rechten Seite von (5) nur Summanden auftreten, die auch auf der rechten Seite von (4) auftreten. Summieren wir (5) über k, dann treten in der Gesamtsumme auch alle Summanden auf, die in (4) auftreten. Damit gilt

$$\mu(A) = \sum_k \mu(B_k),$$

was zu zeigen war.

Auf diese Weise ergibt sich insbesondere die Additivität der Elementarmaße im n-dimensionalen euklidischen Raum aus der Additivität des Maßes auf der Geraden.

Ein Maß (2), das durch die Formel (3) auf einem Semiring (1) definiert wird, heißt *Produkt der Maße* $\mu_1, \mu_2, \ldots, \mu_n$ oder *Produktmaß.*

Satz 2. *Sind $\mu_1, \mu_2, \ldots, \mu_n$ σ-additive Maße, dann ist auch das Maß*

$$\mu = \mu_1 \times \mu_2 \times \cdots \times \mu_n$$

σ-additiv.

Beweis. Wir beweisen den Satz für $n = 2$. Es sei $C \in \mathfrak{S}_1 \times \mathfrak{S}_2$, $C = \bigcup_{n=1}^{\infty} C_n$, $C_n \in \mathfrak{S}_1 \times \mathfrak{S}_2$, $C_i \cap C_j = \emptyset$ $(i \neq j)$ und

$$C = A \times B, \quad A \in \mathfrak{S}_1, \quad B \in \mathfrak{S}_2,$$

$$C_n = A_n \times B_n, \quad A_n \in \mathfrak{S}_1, \quad B_n \in \mathfrak{S}_2.$$

Ist X der Raum, in dem alle Mengen $A, A_1, A_2, \ldots$ liegen, dann setzen wir auf X

$$f_n(x) = \begin{cases} \mu_2(B_n), & \text{falls } x \in A_n, \\ 0, & \text{falls } x \notin A_n. \end{cases}$$

Man sieht sofort, daß für $x \in A$

$$\sum_n f_n(x) = \mu_2(B)$$

ist. Für das Integral bezüglich der Lebesgueschen Fortsetzung λ_1 des Maßes μ_1 gilt einerseits nach dem Satz von B. LEVI (Satz 7 aus 5.5.)

$$\sum_n \int_A f_n(x)\,d\lambda_1 = \int_A \mu_2(B)\,d\lambda_1 = \lambda_1(A)\cdot\mu_2(B) = \mu_1(A)\cdot\mu_2(B) = \mu(C),$$

andererseits ist

$$\int_A f_n(x)\,d\lambda_1 = \mu_2(B_n)\,\mu_1(A_n) = \mu(C_n),$$

d. h.

$$\sum_n \mu(C_n) = \mu(C),$$

was zu zeigen war.

Sind $\mu_1, \ldots, \mu_n$ σ-additive Maße auf den σ-Algebren $\mathfrak{S}_1, \ldots, \mathfrak{S}_n$, dann bezeichnen wir die Lebesguesche Fortsetzung des Maßes $\mu_1 \times \mu_2 \times \cdots \times \mu_n$ als ihr Produkt und schreiben dafür

$$\mu_1 \otimes \mu_2 \otimes \cdots \otimes \mu_n \quad \text{oder} \quad \bigotimes_k \mu_k.$$

Insbesondere ergibt sich für

$$\mu_1 = \mu_2 = \cdots = \mu_n = \mu$$

die n-te Potenz des Maßes μ:

$$\mu^n = \bigotimes_k \mu_k, \quad \mu_k = \mu.$$

So ist z. B. das n-dimensionale Lebesguesche Maß μ^n die n-te Potenz des Lebesgueschen Maßes μ auf der Geraden.

Zum Abschluß bemerken wir noch, daß das Produktmaß stets vollständig ist (sogar dann, wenn die Maße $\mu_1, \ldots, \mu_n$ nicht vollständig sind).

5.6.3. Darstellung des ebenen Maßes als Integral von Schnitten und die geometrische Bedeutung des Lebesgueschen Integrals. Ist G ein Gebiet in der x, y-Ebene, das durch vertikalen Geraden $x = a$, $x = b$ und die Kurven $y = \varphi(x)$, $y = \psi(x)$ begrenzt wird, so wird bekanntlich sein Flächeninhalt durch das Integral

$$V(G) = \int_a^b \{\varphi(x) - \psi(x)\}\,dx$$

beschrieben. Die Differenz $\varphi(x_0) - \psi(x_0)$ ist dabei gleich der Länge des Schnittes des Gebietes G mit der Vertikalen $x = x_0$. Im folgenden werden wir diese Methode der Flächenmessung auf beliebige Produktmaße

$$\mu = \mu_x \otimes \mu_y$$

übertragen. Dazu setzen wir voraus, daß die Maße μ_x und μ_y auf σ-Algebren definiert, σ-additiv und vollständig (d. h., wenn $B \subset A$ und $\mu(A) = 0$ ist, dann ist B meßbar) sind. Früher haben wir gezeigt, daß solche Maße eine Lebesguesche Fortsetzung besitzen.

Im weiteren sei

$$A \subset X \times Y,$$
$$A_x = \{y: (x, y) \in A\} \qquad (x \text{ fest}),$$
$$A_y = \{x: (x, y) \in A\} \qquad (y \text{ fest}),$$

d. h., wenn X und Y Zahlengeraden sind ($X \times Y$ die Ebene), dann bezeichnet A_{x_0} die Projektion des Schnittes der Menge A mit der vertikalen Geraden $x = x_0$ auf die y-Achse.

Satz 3. *Unter den oben angegebenen Voraussetzungen gilt für eine beliebige μ-meßbare Menge*[1] A

$$\mu(A) = \int_X \mu_y(A_x)\, d\mu_x = \int_Y \mu_x(A_y)\, d\mu_y.$$

Beweis. Es genügt, die Gleichung

$$\mu(A) = \int_X \varphi_A(x)\, d\mu_x, \qquad \varphi_A(x) = \mu_y(A_x), \tag{6}$$

zu beweisen, da die zweite Behauptung des Satzes völlig analog zur ersten ist. Dabei ist zu beachten, daß diese Gleichung auch die Behauptung enthält, daß die Menge A_x für fast alle x (im Sinne des Maßes μ_x) meßbar bezüglich des Maßes μ_y und die Funktion $\varphi_A(x)$ meßbar bezüglich des Maßes μ_x ist. Ohne diese Forderungen hätte die Gleichung (6) keinen Sinn.

Es sei μ die Lebesguesche Fortsetzung des Maßes

$$m = \mu_x \times \mu_y,$$

das auf einem System $\mathfrak{S}_m$ von Mengen der Form

$$A = A_{y_0} \times A_{x_0}$$

gegeben ist. Für solche Mengen gilt (6) offensichtlich, da dann

$$\varphi_A(x) = \begin{cases} \mu_y(A_{x_0}) & \text{für} \quad x \in A_{y_0}, \\ 0 & \text{für} \quad x \notin A_{y_0} \end{cases}$$

ist. Ohne Mühe erhält man damit (6) auch für die Mengen aus dem Ring $\mathfrak{R}(\mathfrak{S}_m)$, die als endliche Vereinigung paarweise disjunkter Mengen aus $\mathfrak{S}_m$ dargestellt werden können.

---

[1]) Die Integration über X ist faktisch eine Integration über $\bigcup_y A_y \subset X$, da außerhalb dieser Menge der Integrand gleich 0 ist. Analog ist $\int_Y = \int_{\bigcup_x A_x}$.

Der Beweis der Gleichung (6) im allgemeinen Fall stützt sich auf folgendes Lemma, dessen Aussage in der Theorie der Lebesgueschen Fortsetzungen auch für sich von Interesse ist.

Lemma. *Für jede μ-meßbare Menge A existiert eine Menge B der Form*

$$B = \bigcap_n B_n, \qquad B_1 \supset B_2 \supset \cdots \supset B_n \supset \cdots,$$

$$B_n = \bigcup_k B_{nk}, \qquad B_{n1} \subset B_{n2} \subset \cdots \subset B_{nk} \subset \cdots, B_{nk} \in \Re(\mathfrak{S}_m),$$

so daß $A \subset B$ und

$$\mu(A) = \mu(B) \tag{7}$$

ist.

Beweis. Nach Definition der Meßbarkeit kann man die Menge A für jedes n so durch eine Menge $C_n = \bigcup_r \Delta_{nr}$, $\Delta_{nr} \in \mathfrak{S}_m$, überdecken, daß

$$\mu(C_n) < \mu(A) + \frac{1}{n}$$

ist. Setzen wir $B_n = \bigcap_{k=1}^{n} C_k$, so haben diese Mengen die Form $B_n = \bigcup_k \delta_{ns}$, $\delta_{ns} \in \mathfrak{S}_m$, und die endlichen Vereinigungen $B_{nk} = \bigcup_{s=1}^{k} \delta_{ns}$ bilden ein System mit den geforderten Eigenschaften. Damit ist das Lemma bewiesen.

Der allgemeine Beweis von (6) ergibt sich nun wie folgt. Auf Grund der Beziehungen

$$\varphi_{B_n}(x) = \lim_{k\to\infty} \varphi_{B_{nk}}(x), \qquad \varphi_{B_{n1}} \leqq \varphi_{B_{n2}} \leqq \cdots,$$

$$\varphi_B(x) = \lim_{n\to\infty} \varphi_{B_n}(x), \qquad \varphi_{B_1} \geqq \varphi_{B_2} \geqq \cdots,$$

überträgt sich nach dem Satz von B. Levi (Satz 7 aus 5.5.) die Gleichheit (6) von den Mengen $B_{nk} \in \Re(\mathfrak{S}_m)$ auf die Mengen B_n und die Menge B. Wenn $\mu(A) = 0$ ist, dann ist auch $\mu(B) = 0$ und

$$\varphi_B(x) = \mu_y(B_x) = 0.$$

Wegen $A_x \subset B_x$ ergibt sich daraus die Meßbarkeit von A_x für fast alle x und

$$\varphi_A(x) = \mu_y(A_x) = 0,$$

$$\int \varphi_A(x)\, d\mu_x = 0 = \mu(A),$$

d. h., die Formel (6) ist richtig für eine Menge A mit dem Maß Null. Im allgemeinen Fall stellen wir A in der Form $B \setminus C$ dar, wobei gemäß (7)

$$\mu(C) = 0$$

ist. Aus der Richtigkeit von (6) für die Mengen B und C folgt nun, wie man leicht sieht, auch die Gültigkeit dieser Formel für die Menge A selbst. Damit ist der Beweis von Satz 3 vollständig.

Ist Y die Zahlengerade, μ_y das auf ihr definierte Lebesguesche Maß und A eine Punktmenge der Form

$$\{(x, y): x \in M, 0 \leqq y \leqq f(x)\}, \tag{8}$$

wobei M eine beliebige μ_x-meßbare Menge und $f(x)$ eine nichtnegative Funktion ist, so ergibt sich

$$\mu_y(A_x) = \begin{cases} f(x) & \text{für} \quad x \in M, \\ 0 & \text{für} \quad x \notin M, \end{cases}$$

$$\mu(A) = \int_M f(x)\, d\mu_x,$$

d. h., es gilt folgender Satz.

Satz 4. *Das Lebesguesche Integral einer nichtnegativen Funktion $f(x)$ über die Menge M ist gleich dem Maß $\mu = \mu_x \times \mu_y$ der durch* (8) *definierten Menge A.*

Wenn X die Zahlengerade, M ein Intervall und $f(x)$ eine nach Riemann integrierbare Funktion ist, dann liefert dieser Satz die bekannte Interpretation des Integrals als Flächeninhalt der zwischen der Funktionskurve und x-Achse gelegenen Fläche.

5.6.4. Der Satz von Fubini. Im weiteren werden wir ein dreifaches Mengenprodukt $U = X \times Y \times Z$ betrachten, wobei auf den Mengen X, Y, Z die Maße μ_x, μ_y, μ_z gegeben sein sollen. Wie man leicht nachprüft, kann das Maß

$$\mu_u = \mu_x \otimes \mu_y \otimes \mu_z$$

auf U durch

$$(\mu_x \otimes \mu_y) \otimes \mu_z$$

oder durch

$$\mu_x \otimes (\mu_y \otimes \mu_z)$$

definiert werden. Beide Definitionen führen zu demselben Maß.

Grundlegend für die Theorie der mehrfachen Integrale ist der folgende Satz.

Satz 5 (Fubini). *Sind μ_x, μ_y σ-additive, vollständige, auf σ-Algebren definierte Maße, ist*

$$\mu = \mu_x \otimes \mu_y$$

und ist die Funktion $f(x, y)$ auf der Menge

$$A \subset X \times Y \tag{9}$$

bezüglich des Maßes μ integrierbar, dann gilt[1])

$$\int_A f(x, y)\, d\mu = \int_X \left(\int_{A_x} f(x, y)\, d\mu_y \right) d\mu_x = \int_Y \left(\int_{A_y} f(x, y)\, d\mu_x \right) d\mu_y. \tag{10}$$

[1]) Vgl. die Fußnote auf S. 313.

Die Behauptung des Satzes schließt die Existenz der inneren Integrale für fast alle Werte der Veränderlichen (über die anschließend integriert wird) ein.

Beweis. Wir führen den Beweis zuerst für den Fall $f(x, y) \geqq 0$. Es sei

$$U = X \times Y \times Z$$

das Produkt zweier beliebiger Mengen X, Y mit der Zahlengeraden Z,

$$\lambda = \mu_x \otimes \mu_y \otimes \mu^1 = \mu \otimes \mu^1$$

das Produktmaß der auf X bzw. Y definierten Maße μ_x, μ_y mit dem Lebesgueschen Maß μ^1 auf Z. Ist

$$W = \{(x, y, z): (x, y) \in A, 0 \leqq z \leqq f(x, y)\},$$

dann ergibt sich einerseits nach Satz 4

$$\lambda(W) = \int_A f(x, y)\, d\mu \tag{11}$$

und andererseits nach Satz 3

$$\lambda(W) = \int_X \xi(W_x)\, d\mu_x, \tag{12}$$

wobei $\xi = \mu_y \times \mu^1$ und $W_x = \{(y, z) : (x, y, z) \in W\}$ ist. Wenden wir auf das Maß ξ nochmals Satz 4 an, so ist

$$\xi(W_x) = \int_{A_x} f(x, y)\, d\mu_y. \tag{13}$$

Aus (11), (12) und (13) folgt

$$\int_A f(x, y)\, d\mu = \int_X \left(\int_{A_x} f(x, y)\, d\mu_y \right) d\mu_x,$$

was zu zeigen war.

Den allgemeinen Fall kann man auf den betrachteten Spezialfall mit Hilfe der Gleichungen

$$f(x, y) = f^+(x, y) - f^-(x, y),$$

$$f^+(x, y) = \frac{|f(x, y)| + f(x, y)}{2}, \quad f^-(x, y) = \frac{|f(x, y)| - f(x, y)}{2}$$

zurückführen.

Bemerkung. Wie die unten angeführten Beispiele zeigen, zieht die Existenz der iterierten Integrale

$$\int_X \left(\int_{A_x} f\, d\mu_y \right) d\mu_x \quad \text{und} \quad \int_Y \left(\int_{A_y} f\, d\mu_x \right) d\mu_y \tag{14}$$

im allgemeinen nicht die Gleichheit (10) und nicht die Integrierbarkeit der Funktion $f(x, y)$ auf A nach sich. *Existiert jedoch wenigstens eines der beiden Integrale*

$$\int_X \left(\int_{A_x} |f(x, y)| \, d\mu_y \right) d\mu_x \quad \textit{oder} \quad \int_Y \left(\int_{A_y} |f(x, y)| \, d\mu_x \right) d\mu_y, \tag{15}$$

dann ist $f(x, y)$ auf A integrierbar und die Gleichung (10) *richtig.*

Wenn z. B. das erste Integral von (15) existiert und gleich M ist, dann ist die Funktion $f_n(x, y) = \min \{|f(x, y|, n\}$ meßbar, beschränkt und folglich summierbar auf A. Nach dem Satz von FUBINI ist

$$\int_A f_n(x, y) \, d\mu = \int_X \left(\int_{A_x} f_n(x, y) \, d\mu_y \right) d\mu_x \leqq M. \tag{16}$$

Da die Funktionen f_n eine nichtfallende, fast überall gegen $|f(x, y)|$ konvergente Folge bilden, folgt aus (16) nach dem Satz von B. LEVI, daß $|f(x, y)|$ auf A integrierbar ist. Damit ist auch $f(x, y)$ auf A integrierbar. Aus dem Satz von FUBINI ergibt sich nun unsere Behauptung.

Wir haben den Satz von FUBINI unter der Voraussetzung bewiesen, daß die Maße μ_x und μ_y (und damit auch μ) endlich sind. Der Satz ist jedoch auch im Fall σ-endlicher Maße richtig (vgl. etwa [12], S. 208).

Zum Schluß geben wir noch zwei Beispiele von Funktionen an, für die die iterierten Integrale (14) existieren, die Gleichung (10) jedoch nicht erfüllt ist.

1. Es sei

$$A = [-1, 1]^2$$

und

$$f(x, y) = \frac{xy}{(x^2 + y^2)^2}.$$

Dann ist

$$\int_{-1}^{1} f(x, y) \, dx = 0 \qquad \text{für } y \neq 0$$

und

$$\int_{-1}^{1} f(x, y) \, dy = 0 \qquad \text{für } x \neq 0.$$

Daher gilt

$$\int_{-1}^{1} \left(\int_{-1}^{1} f(x, y) \, dx \right) dy = \int_{-1}^{1} \left(\int_{-1}^{1} f(x, y) \, dy \right) dx = 0,$$

während das zweifache Lebesguesche Integral über das Quadrat nicht existiert, denn es ist

$$\int_{-1}^{1} \int_{-1}^{1} |f(x, y)| \, dx \, dy \geqq \int_0^1 dr \int_0^{2\pi} \frac{|\sin \varphi \cos \varphi|}{r} \, d\varphi = 2 \int_0^1 \frac{dr}{r} = \infty.$$

2. Es sei

$$A = [0, 1]^2$$

und

$$f(x, y) = \begin{cases} 0 & \text{für } x = 0,\ y = 0, \\ \dfrac{x^2 - y^2}{(x^2 + y^2)^2} & \text{sonst.} \end{cases}$$

Dann ist

$$\int_0^1 \left(\int_0^1 f(x, y)\, dy \right) dx = \frac{\pi}{4}$$

und

$$\int_0^1 \left(\int_0^1 f(x, y)\, dx \right) dy = -\frac{\pi}{4}.$$

Es gilt nämlich für $x \neq 0$

$$f(x, y) = \frac{\partial}{\partial y} \left(\frac{y}{x^2 + y^2} \right)$$

und daher

$$\int_0^1 f(x, y)\, dy = \frac{1}{x^2 + 1}.$$

Daraus folgt, daß das erste iterierte Integral den angegebenen Wert hat.
Analog berechnet man das zweite iterierte Integral.

6. Das unbestimmte Lebesguesche Integral. Theorie der Differentiation

In diesem Kapitel werden wir das Lebesguesche Integral reeller Funktionen über Teilmengen der Zahlengeraden bezüglich des üblichen Lebesgueschen Maßes betrachten.

Ist f eine auf der Menge X mit dem Maß μ integrierbare Funktion, dann existiert das Integral

$$\int_A f(x)\,d\mu \tag{*}$$

für jedes meßbare $A \subset X$ und stellt bei festem f eine Mengenfunktion dar, die für alle meßbaren Untermengen $A \subset X$ erklärt ist. Ein solches Integral heißt *unbestimmtes Lebesguesches Integral*. Wird als Menge X ein Intervall der Zahlengeraden und als Maß μ das übliche Lebesguesche Maß auf der Geraden (wir schreiben dann im Integral dt anstelle von $d\mu$) betrachtet, dann ist für ein Teilintervall $A \subset X$ das Integral (*) eine Funktion der Intervallendpunkte. Halten wir einen der Endpunkte des Integrationsintervalls fest, etwa den linken, so können wir das unbestimmte Integral über das Intervall $[a, x]$

$$\int_a^x f(t)\,dt$$

als Funktion einer reellen Veränderlichen x untersuchen. Dieses Problem führt uns zur Betrachtung einiger wichtiger Funktionenklassen auf der Zahlengeraden. Die allgemeine Untersuchung des Lebesgueschen Integrals (von einer festen Funktion f) als Mengenfunktion wird in 6.5. durchgeführt.

Aus der Grundvorlesung über Analysis sind folgende grundlegende Beziehungen über den Zusammenhang zwischen Differentiation und Integration bekannt: Ist f eine stetige Funktion und F eine Funktion mit stetiger Ableitung, dann gilt

a) $$\frac{d}{dx}\int_a^x f(t)\,dt = f(x),$$

b) $$\int_a^b F'(t)\,dt = F(b) - F(a).$$

Nach der Einführung des Lebesgueschen Integrals entsteht die Frage, ob die Gleichung a) auch für Lebesgue-integrierbare Funktionen richtig ist und für welche (möglichst umfangreiche) Funktionenklasse Gleichung b) gilt, wenn man die dort auftretenden Integrale im Lebesgueschen Sinne versteht.

Der Beantwortung dieser Frage sind die folgenden Abschnitte gewidmet.

6.1. Monotone Funktionen. Differentiation des Integrals nach der oberen Grenze

6.1.1. Grundlegende Eigenschaften monotoner Funktionen. Die Untersuchung der Eigenschaften des Lebesgueschen Integrals

$$\Phi(x) = \int_a^x f(t)\, dt \tag{1}$$

als Funktion der oberen Grenze beginnen wir mit folgender offensichtlichen, aber wichtigen Bemerkung: Wenn die Funktion $f(x)$ nichtnegativ ist, dann ist $\Phi(x)$ eine monoton nichtfallende Funktion. Da jede integrierbare Funktion als Differenz zweier nichtnegativer integrierbarer Funktionen

$$f(t) = f_+(t) - f_-(t) \tag{2}$$

dargestellt werden kann, ist das Integral (1) stets in eine Differenz zweier monoton nichtfallender Funktionen zerlegbar. Daher kann die Untersuchung des Lebesgueschen Integrals als Funktion der oberen Grenze auf die Untersuchung von monotonen Funktionen zurückgeführt werden. Das erweist sich als vorteilhaft, da monotone Funktionen (unabhängig von ihrer konkreten Struktur) eine Reihe von einfachen, aber wichtigen Eigenschaften besitzen. Wir stellen nun diese Eigenschaften dar. Dazu erinnern wir zunächst an einige Begriffe.

Überall, wo nicht ausdrücklich das Gegenteil vorausgesetzt wird, betrachten wir reellwertige Funktionen, die auf einem Intervall der Zahlengeraden definiert sind.

Eine Funktion f heißt *monoton nichtfallend*, wenn aus $x_1 \leqq x_2$

$$f(x_1) \leqq f(x_2)$$

folgt. Entsprechend ist die Bezeichnung „monoton nichtwachsend" zu verstehen.

Ist f eine beliebige Funktion auf der Zahlengeraden, dann heißt der Grenzwert[1])

$$\lim_{h \to 0+0} f(x_0 + h)$$

(wenn er existiert) *rechtsseitiger Grenzwert* der Funktion f im Punkt x_0 und wird mit $f(x_0 + 0)$ bezeichnet. Entsprechend ist $f(x_0 - 0)$, der *linksseitige Grenzwert* der Funk-

---

[1]) Unter dem Symbol $h \to 0 + 0$ ist zu verstehen, daß h gegen 0 strebt und dabei nur positive Werte annimmt.

tion f im Punkt x_0, definiert. Die Gleichheit $f(x_0 + 0) = f(x_0 - 0)$ drückt aus, daß die Funktion f im Punkt x_0 entweder stetig ist oder dort eine hebbare Unstetigkeit besitzt. Ein Punkt, in dem diese beiden Grenzwerte existieren, aber nicht gleich sind, heißt *Unstetigkeitspunkt erster Art*, die Differenz $f(x_0 + 0) - f(x_0 - 0)$ wird dann der *Sprung* der Funktion in diesem Punkt genannt

Wenn $f(x_0) = f(x_0 - 0)$ ist, heißt f *von links stetig* (*linksseitig stetig*) im Punkt x_0; Wenn $f(x_0) = f(x_0 + 0)$ ist, heißt f *von rechts stetig* (*rechtsseitig stetig*) in x_0.

Wir beweisen nun einige Grundeigenschaften monotoner Funktionen. Alle Aussagen, die wir hierbei für monoton nichtfallende Funktionen formulieren, sind selbstverständlich auch für monoton nichtwachsende Funktionen (mit entsprechender Modifikation der Formulierung) richtig.

1. *Jede auf $[a, b]$ monoton nichtfallende Funktion f ist meßbar und beschränkt und deshalb integrierbar.*

Nach Definition der Monotonie ist

$$f(a) \leqq f(x) \leqq f(b) \quad \text{auf } [a, b].$$

Damit ist für eine beliebige Konstante c die Menge

$$A_c = \{x : f(x) < c\}$$

entweder ein abgeschlossenes Intervall, ein halboffenes Intervall oder die leere Menge, d. h. in jedem Fall meßbar, was zu zeigen war.

2. *Eine monotone Funktion besitzt nur Unstetigkeitspunkte erster Art.*

Ist x_0 ein beliebiger Punkt aus $[a, b]$ und $x_n \to x_0$; $x_n < x_0$, dann ist die Folge $\{f(x_n)\}$ beschränkt nach unten und oben (z. B. durch $f(a)$ bzw. $f(b)$) und kann folglich nur einen Häufungspunkt besitzen. Für verschiedene solche Folgen $\{x_n\}$ ergibt sich stets der gleiche Häufungspunkt für die Folge $\{f(x_n)\}$ der Funktionswerte, da die Existenz verschiedener Häufungspunkte der Monotonie der Funktion f widerspricht. Daher existiert $f(x_0 - 0)$. Entsprechend ergibt sich die Existenz von $f(x_0 + 0)$.

Eine monotone Funktion ist nicht notwendig stetig. Aber es gilt folgender Satz:

3. *Die Menge der Unstetigkeitspunkte einer monotonen Funktion ist höchstens abzählbar.*

Da die Summe einer beliebigen endlichen Anzahl von Sprüngen einer auf dem Intervall $[a, b]$ monotonen Funktion f die Differenz $f(b) - f(a)$ nicht übersteigen darf, ist für jedes n die Anzahl der Sprünge, die größer als $1/n$ sind, endlich. Numerieren wir diese Sprungstellen für $n = 1, 2, \ldots$ nacheinander, so ergibt sich ihre Abzählbarkeit.

Die einfachsten unter den monotonen Funktionen sind die sogenannten *Sprungfunktionen*. Diese Funktionen werden auf folgende Weise definiert. In endlich oder abzählbar vielen Punkten

$$x_1, x_2, \ldots, x_n, \ldots$$

des Intervalls $[a, b]$ werden positive Zahlen h_n mit $\sum_n h_n < \infty$ vorgegeben und der Funktionswert durch

$$f(x) = \sum_{x_n < x} h_n$$

festgelegt.

Offensichtlich sind solche Funktionen monoton nichtfallend. Außerdem sind sie in jedem Punkt von links stetig[1]), und die Menge ihrer Unstetigkeitspunkte stimmt mit der Menge $\{x_n\}$ überein. Da nämlich

$$f(x - 0) = \lim_{\varepsilon \to 0+0} f(x - \varepsilon) = \lim_{\varepsilon \to 0+0} \sum_{x_n < x - \varepsilon} h_n$$

ist und jedes x_n, das der Bedingung $x_n < x$ genügt, für hinreichend kleines ε auch die Bedingung $x_n < x - \varepsilon$ erfüllt, ist der letzte Grenzwert gleich $\sum_{x_n < x} h_n = f(x)$. Somit ist

$$f(x - 0) = f(x).$$

Wenn der Punkt x mit einem der Punkte x_n übereinstimmt, d. h. $x = x_{n_0}$ ist, dann folgt

$$f(x_{n_0} + 0) = \lim_{\varepsilon \to 0+0} f(x_{n_0} + \varepsilon) = \lim_{\varepsilon \to 0+0} \sum_{x_n < x_{n_0} + \varepsilon} h_n = \sum_{x_n \leqq x_{n_0}} h_n,$$

also

$$f(x_{n_0} + 0) - f(x_{n_0} - 0) = h_{n_0}.$$

Wenn schließlich x mit keinem der Punkte x_n übereinstimmt, dann ist die Sprungfunktion in x stetig (man beweise diesen Sachverhalt!).

Im weiteren werden wir unter einer Sprungfunktion eine beliebige Funktion verstehen, die auf die oben beschriebene Weise erklärt ist. Den einfachsten Typ von Sprungfunktionen stellen Treppenfunktionen dar, bei denen die Menge der Unstetigkeitspunkte $\{x_n\}$ eine monotone Folge bildet,

$$x_1 < x_2 < \cdots < x_n < \cdots.$$

Im allgemeinen Fall kann eine Sprungfunktion jedoch eine kompliziertere Struktur haben. Zum Beispiel, wenn $\{x_n\}$ die Menge aller rationalen Zahlen des Intervalls $[a, b]$ und $h_n = (1/2)^n$ ist, dann wird durch (3) eine Sprungfunktion definiert, die in allen rationalen Punkten unstetig und in allen irrationalen Punkten stetig ist.

---

[1]) Wenn wir f durch die Formel

$$f(x) = \sum_{x_n \leqq x} h_n$$

definiert hätten, so wäre diese Funktion in jedem Punkt *von rechts stetig*. Im folgenden werden wir jedoch, wenn nicht ausdrücklich das Gegenteil vorausgesetzt wird, alle betrachteten monotonen Funktionen als von links stetig voraussetzen.

Einen anderen Typ von monotonen Funktionen, in gewissem Sinn das Gegenstück zu den Sprungfunktionen, stellen die stetigen monotonen Funktionen dar. Es gilt nämlich der folgende Satz.

4. *Jede von links stetige monotone Funktion kann als Summe aus einer stetigen monotonen Funktion und einer (von links stetigen) Sprungfunktion dargestellt werden. Diese Darstellung ist eindeutig.*

Es sei f eine von links stetige monoton nichtfallende Funktion mit den Unstetigkeitspunkten $x_1, x_2, \ldots$ und den Sprüngen $h_1, h_2, \ldots$ in diesen Punkten. Setzen wir

$$H(x) = \sum_{x_n < x} h_n ,$$

dann ist die Differenz

$$\varphi = f - H$$

eine stetige nichtfallende Funktion. Um dies zu beweisen, betrachten wir die Differenz

$$\varphi(x'') - \varphi(x') = [f(x'') - f(x')] - [H(x'') - H(x')],$$

wobei $x' < x''$ sei. Auf der rechten Seite dieser Gleichung steht die Differenz zwischen dem vollen Zuwachs der Funktion f auf dem Intervall $[x', x'']$ und der Summe ihrer Sprünge auf diesem Intervall. Es ist klar, daß diese Größe nichtnegativ ist, d. h., φ ist eine monoton nichtfallende Funktion. Desweiteren gilt für einen beliebigen Punkt x^*

$$\varphi(x^* - 0) = f(x^* - 0) - H(x^* - 0) = f(x^* - 0) - \sum_{x_n < x^*} h_n ,$$

$$\varphi(x^* + 0) = f(x^* + 0) - H(x^* + 0) = f(x^* + 0) - \sum_{x_n \leqq x^*} h_n .$$

Daraus folgt

$$\varphi(x^* + 0) - \varphi(x^* - 0) = f(x^* + 0) - f(x^* - 0) - h^* = 0$$

(h^* ist der Sprung der Funktion H in x^*), woraus auf Grund der Stetigkeit von links der Funktionen f und H die Stetigkeit der Funktion φ folgt, was zu zeigen war.

6.1.2. Differenzierbarkeit der monotonen Funktionen. Wir wenden uns jetzt der Frage zu, ob eine monotone Funktion differenzierbar ist oder nicht. Der folgende Satz faßt die Antwort zusammen.

Satz 1 (Lebesgue). *Eine auf $[a, b]$ monotone Funktion f besitzt fast überall in diesem Intervall eine endliche Ableitung.*

Bevor wir diese Aussage beweisen, führen wir einige dazu notwendige Begriffe ein. Bekanntlich versteht man unter der Ableitung einer Funktion f im Punkt x_0 den

Grenzwert von

$$\frac{f(x) - f(x_0)}{x - x_0} \tag{4}$$

für $x \to x_0$. Unabhängig davon, ob dieser Grenzwert endlich ist oder nicht existiert, besitzen die folgenden Häufungswerte einen Sinn (sie können auch unendlich sein):

$$\Lambda_r = \overline{\lim_{x \to x_0+0}} \frac{f(x) - f(x_0)}{x - x_0} \quad \text{(obere rechte Ableitungszahl)},$$

$$\lambda_r = \underline{\lim_{x \to x_0+0}} \frac{f(x) - f(x_0)}{x - x_0} \quad \text{(untere rechte Ableitungszahl)},$$

$$\Lambda_l = \overline{\lim_{x \to x_0-0}} \frac{f(x) - f(x_0)}{x - x_0} \quad \text{(obere linke Ableitungszahl)},$$

$$\lambda_l = \underline{\lim_{x \to x_0-0}} \frac{f(x) - f(x_0)}{x - x_0} \quad \text{(untere linke Ableitungszahl)},$$

Abb. 19 gibt eine Interpretation dieser Zahlen als Anstiege gewisser „Grenzsekanten". Entsprechend der Definition ist

$$\lambda_r \leqq \Lambda_r \quad \text{und} \quad \lambda_l \leqq \Lambda_l.$$

Wenn Λ_r und λ_r endlich sind und übereinstimmen, dann ist ihr gemeinsamer Wert die rechtsseitige Ableitung der Funktion $f(x)$ im Punkt x_0. Analog ist im Fall $\Lambda_l = \lambda_l$

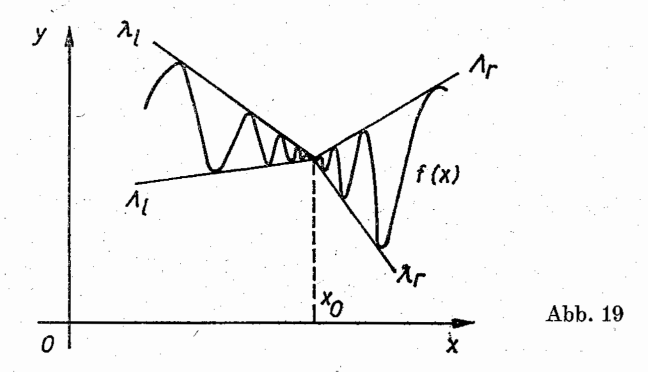


Abb. 19

diese Zahl gleich der linksseitigen Ableitung von f in x_0. Die Existenz einer endlichen Ableitung der Funktion f im Punkt x_0 ist gleichbedeutend damit, daß alle Ableitungszahlen in diesem Punkt endlich sind und übereinstimmen. Daher kann man die Aussage des Satzes von Lebesgue in der folgenden Weise formulieren: *Für eine auf* $[a, b]$ *monotone Funktion ist fast überall auf* $[a, b]$

$$-\infty < \lambda_l = \lambda_r = \Lambda_l = \Lambda_r < \infty$$

erfüllt.

Aufgabe. Es sei $f^*(x) = -f(x)$. Wie sind die Ableitungszahlen von f^* und f miteinander verknüpft?

Man beantworte dieselbe Frage auch für den Übergang von $f(x)$ zu $f(-x)$.

Der Beweis des Satzes von LEBESGUE stützt sich auf das folgende Lemma, das auch im weiteren noch benutzt wird. Zu seiner Formulierung verwenden wir folgende Sprechweise. Für eine auf $[a, b]$ stetige Funktion $g(x)$ heißt ein Punkt x_0 *unsichtbar von rechts*, wenn es einen Punkt $\xi \in (x_0, b]$ gibt, für den $g(x_0) < g(\xi)$ ist (siehe Abb. 20).

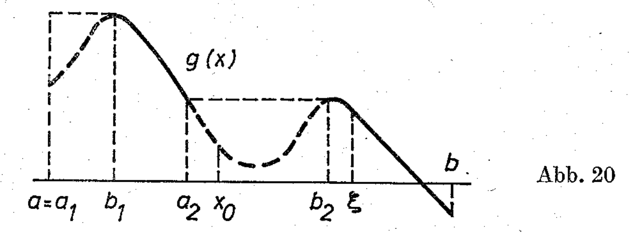


Abb. 20

Lemma (F. RIESZ). *Für eine beliebige stetige Funktion g ist die Menge der von rechts unsichtbaren Punkte eine offene Untermenge des Intervalls $[a, b]$*[1]) *und folglich als Vereinigung endlich oder abzählbar vieler paarweise disjunkter Intervalle (a_k, b_k) darstellbar (eventuell tritt dabei auch das halboffene Intervall $[a_1, b_1)$ mit $a_1 = a$ auf). In den Endpunkten dieser Intervalle gilt*

$$g(a_k) \leqq g(b_k). \tag{5}$$

Beweis des Lemmas. Wenn x_0 ein von rechts unsichtbarer Punkt für g ist, dann besitzt ein beliebiger anderer Punkt aus einer hinreichend kleinen Umgebung von x_0 (rechtsseitige Umgebung für $x_0 = a$) auf Grund der Stetigkeit von g dieselbe Eigenschaft. Daher ist die Menge aller solcher Punkte offen in $[a, b]$.

Ist (a_k, b_k) ein Intervall aus der Darstellung der offenen Menge der unsichtbaren Punkte und nehmen wir

$$g(a_k) > g(b_k) \tag{6}$$

an, so gibt es im Intervall (a_k, b_k) einen inneren Punkt x_0, für den auch $g(x_0) > g(b_k)$ ist. Ist x^* der am weitesten rechts gelegene der Punkte $x \in (a_k, b_k)$ mit $g(x) \geqq g(x_0)$ dann existiert ein Punkt $\xi \in (x^*, b]$, für den $g(\xi) > g(x^*)$ ist. Der Punkt ξ kann nun nach Definition des Punktes x^* und $g(x^*) > g(b_k)$ nicht im Intervall (a_k, b_k) liegen. Andererseits kann er auch nicht rechts von b_k liegen, weil dann aus $g(b_k) < g(x_0) < g(\xi)$ folgen würde, daß b_k entgegen der Voraussetzung ein von rechts unsichtbarer Punkt wäre. Dieser Widerspruch zeigt, daß die Ungleichung (6) falsch sein muß, d. h., $g(a_k) \leqq g(b_k)$ gilt, was zum Beweis des Lemmas zu zeigen war. Der Leser prüft ohne Mühe nach, daß sogar $g(a_k) = g(b_k)$ gilt, wenn $a_k \neq a$ ist.

[1]) Anm. d. Übers.: Die offenen Mengen auf $[a, b]$ sind die Durchschnitte der offenen Mengen auf $(-\infty, \infty)$ mit $[a, b]$. Insbesondere sind also $[a, c)$, $(c, b]$ und $[a, b]$ mit $c \in (a, b)$ offene Mengen.

Bemerkung. Ein Punkt $x_0 \in [a, b]$ heißt *von links unsichtbar* für die stetige Funktion $g(x)$, wenn es einen Punkt $\xi \in [a, x_0)$ gibt, für den $g(\xi) > g(x_0)$ ist. Dieselben Überlegungen wie oben zeigen, daß die Menge aller von links unsichtbaren Punkte eine Vereinigung aus endlich oder abzählbar vielen paarweise disjunkten Intervallen (a_k, b_k) ist, wobei

$$g(a_k) \geqq g(b_k)$$

gilt.

Wir gehen nun zum Beweis des Satzes von Lebesgue über. Zuerst zeigen wir diesen Satz für den Fall, daß f eine stetige monoton nichtfallende Funktion ist. Dazu genügt es zu beweisen, daß fast überall auf $[a, b]$

a) $\Lambda_r < \infty$,

b) $\lambda_l \geqq \Lambda_r$

ist. Setzen wir nämlich $f^*(x) = -f(2x_0 - x)$ in einer Umgebung von x_0, so ist f^* ebenfalls eine stetige monoton nichtfallende Funktion, und ihre Ableitungszahlen Λ_r^*, λ_l^* stehen mit denen der Funktion f (jeweils auf x_0 bezogen), wie man leicht nachprüft, in folgendem Zusammenhang:

$$\Lambda_r^* = \Lambda_l, \qquad \lambda_l^* = \lambda_r$$

(vgl. die Aufgabe auf S. 325). Schreiben wir b) für f auf, so folgt aus diesen Beziehungen

$$\lambda_r \geqq \Lambda_l, \tag{7}$$

und wenn wir b) und (7) zusammen mit den offensichtlichen Ungleichungen $\lambda_l \leqq \Lambda_l$, $\lambda_r \leqq \Lambda_r$ fortlaufend aufschreiben, gelangen wir zu

$$\Lambda_r \leqq \lambda_l \leqq \Lambda_l \leqq \lambda_r \leqq \Lambda_r.$$

Daraus folgt

$$\lambda_l = \lambda_r = \Lambda_l = \Lambda_r.$$

Wir zeigen nun, daß fast überall $\Lambda_r < \infty$ ist. Nehmen wir an, daß $\Lambda_r = \infty$ in einem Punkt x_0 ist, dann gibt es für eine beliebige Konstante $C > 0$ stets rechts von x_0 einen Punkt ξ, so daß

$$\frac{f(\xi) - f(x_0)}{\xi - x_0} > C.$$

ist, d. h.

$$f(\xi) - f(x_0) > C(\xi - x_0)$$

oder

$$f(\xi) - C\xi > f(x_0) - Cx_0.$$

Daraus folgt, daß x_0 ein von rechts unsichtbarer Punkt für die Funktion

$$g(x) = f(x) - Cx$$

ist. Nach dem Lemma von F. Riesz ist die Menge aller solcher Punkte offen, und es gilt

$$f(a_k) - Ca_k \leqq f(b_k) - Cb_k$$

für die Endpunkte a_k, b_k der Intervalle (a_k, b_k) aus der Darstellung dieser offenen Mengen als Vereinigung paarweise disjunkter Intervalle. Schreiben wir die letzte Ungleichung in der Form

$$(b_k - a_k) \leqq \frac{f(b_k) - f(a_k)}{C}$$

und summieren über k, so ergibt sich

$$\sum_k (b_k - a_k) \leqq \sum_k \frac{f(b_k) - f(a_k)}{C} \leqq \frac{f(b) - f(a)}{C}.$$

Da C beliebig groß gewählt werden kann, folgt daraus, daß die Menge aller Punkte, in denen $\Lambda_r = \infty$ ist, durch Intervalle (a_k, b_k) überdeckt werden kann, deren Gesamtlänge beliebig klein gemacht werden kann, d. h., diese Menge hat das Maß Null.

Auf ähnliche Weise wird nun unter (zweimaliger) Verwendung des Rieszschen Lemmas gezeigt, daß fast überall $\lambda_l \geqq \Lambda_r$ gilt. Diese Aussage ist gleichbedeutend damit, daß $\lambda_l < \Lambda_r$ nur auf einer Menge M vom Maß Null ist. Da diese Menge M sich ihrerseits als abzählbare Vereinigung von Mengen $E_{c,C}$ der Punkte x darstellen läßt, in denen $\lambda_l < c$ und $\Lambda_r > C$ für feste rationale Zahlen c und C $(0 < c < C < \infty)$ ist, genügt es, $\mu(E_{c,C}) = 0$ für beliebige c und C zu zeigen. Dafür beweisen wir zunächst:

Für ein beliebiges Intervall $(\alpha, \beta) \subset [a, b]$ *ist*

$$\mu\big(E_{c,C} \cap (\alpha, \beta)\big) \leqq \varrho(\beta - \alpha), \qquad \varrho = c/C.$$

Betrachtet man die Menge derjenigen $x \in (\alpha, \beta)$, für die $\lambda_l < c$ ist, dann gibt es für jeden Punkt x dieser Menge ein $\xi < x$, so daß

$$\frac{f(\xi) - f(x)}{\xi - x} < c,$$

ist, d. h.

$$f(\xi) - c\xi > f(x) - cx,$$

oder anders ausgedrückt: x ist ein von links unsichtbarer Punkt für die Funktion $f(x) - cx$. Nach dem Lemma von F. Riesz (vgl. die Bemerkung auf S. 326) ist die Menge solcher x als Vereinigung von höchstens abzählbar vielen Intervallen (α_k, β_k)

aus (α, β) darstellbar, wobei

$$f(\alpha_k) - c\alpha_k \geqq f(\beta_k) - c\beta_k,$$

d. h.

$$f(\beta_k) - f(\alpha_k) \leqq c(\beta_k - \alpha_k) \tag{8}$$

ist. Betrachten wir nun in jedem der Intervalle (α_k, β_k) die Menge G_k derjenigen x, für die $\Lambda_r > C$ ist, so erhalten wir wieder unter Verwendung des Rieszschen Lemmas, daß G_k als Vereinigung höchstens abzählbar vieler paarweise disjunkter Intervalle $(\alpha_{kj}, \beta_{kj})$ darstellbar ist und dabei

$$\beta_{kj} - \alpha_{kj} \leqq \frac{1}{C} [f(\beta_{kj}) - f(\alpha_{kj})] \tag{9}$$

gilt. Da die Menge $E_{c,C} \cap (\alpha, \beta)$ offensichtlich durch das System der Intervalle $(\alpha_{kj}, \beta_{kj})$ $(k, j = 1, 2, \ldots)$ überdeckt wird und nach (8) und (9)

$$\begin{aligned}\sum_{k,j} (\beta_{kj} - \alpha_{kj}) &\leqq \frac{1}{C} \sum_{k,j} [f(\beta_{kj}) - f(\alpha_{kj})] \\ &\leqq \frac{1}{C} \sum_k [f(\beta_k) - f(\alpha_k)] \leqq \frac{c}{C} \sum_k (\beta_k - \alpha_k) \\ &\leqq \varrho(\beta - \alpha)\end{aligned}$$

ist, ergibt sich die obige Abschätzung für das Maß der Menge $E_{c,C} \cap (\alpha, \beta)$.

Auf der Grundlage dieser Abschätzung ergibt sich nun aus dem nächsten Lemma sofort, daß $\mu(E_{c,C}) = 0$ ist.

Lemma. *Ist A eine meßbare Untermenge des Intervalls $[a, b]$ und ist für ein beliebiges Intervall $(\alpha, \beta) \subset [a, b]$ die Abschätzung $\mu(A \cap (\alpha, \beta)) \leqq \varrho(\beta - \alpha)$ erfüllt, wobei $0 < \varrho < 1$ ist, dann ist $\mu(A) = 0$.*

Beweis. Es sei $\mu(A) = t$. Dann existiert für jedes $\varepsilon > 0$ eine offene Menge $G = \bigcup_m (a_m, b_m)$ ((a_m, b_m) paarweise disjunkte Intervalle), so daß $A \subset G$ und $\sum_m (b_m - a_m) < t + \varepsilon$ ist. Setzen wir $t_m = \mu(A \cap (a_m, b_m))$, so ist $t = \sum t_m$. Nach der Voraussetzung des Lemmas ist andererseits $t \leqq \varrho \sum_m (b_m - a_m) < \varrho(t + \varepsilon)$, woraus wegen $0 < \varrho < 1$ die Aussage $t = 0$ folgt.

Damit ist das letzte Lemma bewiesen und gleichzeitig der Beweis des Satzes von Lebesgue für den Fall abgeschlossen, daß die betrachtete monotone Funktion f stetig ist.

Der Beweis für allgemeinen Fall ergibt sich völlig analog, wenn man die allgemeinere Fassung des Lemmas von Riesz benutzt, die nachfolgend angegeben wird.

Es sei g eine auf dem Intervall $[a, b]$ definierte Funktion, die nur Unstetigkeitsstellen erster Art besitzt. Nennen wir einen Punkt $x_0 \in [a, b]$ *unsichtbar von rechts*

für $g(x)$, wenn ein $\xi \in (x_0, b]$ existiert, so daß

$$\max [g(x_0 - 0), g(x_0), g(x_0 + 0)] < g(\xi)$$

ist, dann ist die Menge der für g von rechts unsichtbaren Punkte offen, und für die Endpunkte a_k, b_k einer Darstellung dieser Menge durch paarweise disjunkte offene Intervalle (a_k, b_k) gilt

$$g(a_k) \leqq g(b_k).$$

Obwohl der Beweis des Satzes 1 ziemlich lang ist, besitzt der Satz selbst einen anschaulichen Sinn. So läßt sich z. B. leicht erklären, daß Λ_r (und Λ_l) notwendig fast überall endlich sein müssen. Der Quotient $\Delta f/\Delta x$, der das Wachsen der Funktion f im Punkt x beschreibt, kann nämlich bei endlichem Gesamtwachstum $f(b) - f(a)$ von f auf $[a, b]$ auf keiner Menge positiven Maßes unendlich sein.

Manchmal erweist sich auch der folgende Satz über die gliedweise Differentiation einer Reihe aus monotonen Funktionen als nützlich.

Satz 2. *Eine auf* $[a, b]$ *überall konvergente Reihe*

$$\sum_{n=1}^{\infty} F_n(x) = F(x) \tag{10}$$

monoton nichtfallender Funktionen $F_n(x)$ *kann fast überall gliedweise definiert werden, d. h., für fast alle Punkte* $x \in [a, b]$ *gilt*

$$\sum_{n=1}^{\infty} F_n'(x) = F'(x).$$

Beweis. Man kann o. B. d. A. voraussetzen, daß alle $F_n(x)$ nichtnegativ und im Punkt $x = a$ gleich 0 sind, da andernfalls das Ersetzen von $F_n(x)$ durch $F_n(x) - F_n(a)$ zu solchen Funktionen führt.

Es sei $E \subset [a, b]$ die Menge aller Punkte x, in denen $F_n'(x)$ (für alle n) und $F'(x)$ existieren. Nach Satz 1 ist $\mu(E) = (b - a)$. Für beliebige Punkte $x \in E$ und $\xi \in [a, b]$ gilt

$$\frac{\sum_{n=1}^{\infty} \{F_n(\xi) - F_n(x)\}}{\xi - x} = \frac{F(\xi) - F(x)}{\xi - x}.$$

Weil $\xi - x$ und $F_n(\xi) - F_n(x)$ auf Grund der Monotonie das gleiche Vorzeichen haben, folgt daraus, daß für jedes N

$$\frac{\sum_{n=1}^{N} \{F_n(\xi) - F_n(x)\}}{\xi - x} \leqq \frac{F(\xi) - F(x)}{\xi - x}.$$

ist. Führt man in dieser Ungleichung den Grenzübergang $\xi \to x$ durch, so ergibt sich

$$\sum_{n=1}^{N} F_n'(x) \leqq F'(x)$$

und daraus wegen $F_n'(x) \geqq 0$ $(n = 1, 2, \ldots)$

$$\sum_{n=1}^{\infty} F_n'(x) \leqq F'(x), \tag{11}$$

d. h., die Reihe der Ableitungen $F_n'(x)$ konvergiert überall auf E. Wir zeigen nun, daß in (11) für fast alle x das Gleichheitszeichen gilt. Dazu bestimmen wir für jedes k eine Partialsumme $S_{n_k}(x)$ der Reihe (10), so daß

$$0 \leqq \big(F(b) - S_{n_k}(b)\big) < \frac{1}{2^k}$$

ist. Da die Funktion $F(x) - S_{n_k}(x) = \sum_{m > n_k} F_m(x)$ monoton nichtfallend ist, gilt für beliebiges x auch

$$0 \leqq F(x) - S_{n_k}(x) < \frac{1}{2^k}.$$

Daher konvergiert die aus nichtfallenden Funktionen bestehende Reihe

$$\sum_{k=1}^{\infty} [F(x) - S_{n_k}(x)] \tag{12}$$

auf dem ganzen Intervall $[a, b]$. Nach dem oben Bewiesenen konvergiert die Reihe

$$\sum_{k=1}^{\infty} [F'(x) - S_{n_k}'(x)], \tag{13}$$

die aus (12) durch gliedweise Differentiation entsteht, fast überall. Da das allgemeine Glied der konvergenten Reihe (13) notwendig gegen 0 strebt, gilt fast überall

$$S_{n_k}'(x) - F'(x) \to 0.$$

Aus dieser Konvergenz einer Teilfolge $\{S_{n_k}'(x)\}$ der überall auf E konvergenten Partialsummenfolge $\{S_n'(x)\}$ von $\sum_{n=1}^{\infty} F_n'(x)$ ergibt sich, daß in (11) fast überall das Gleichheitszeichen stehen muß, d. h. auch auf $[a, b]$ fast überall

$$\sum_{n=1}^{\infty} F_n'(x) = F'(x)$$

ist. Damit ist der Satz bewiesen.

F o l g e r u n g. *Eine Sprungfunktion hat fast überall eine verschwindende Ableitung.*

Das ergibt sich sofort aus Satz 2, wenn die Sprungfunktion als Reihe der nichtfallenden „Stufen"

$$F_n(x) = \begin{cases} 0 & \text{für} \quad x \leqq x_n, \\ h_n & \text{für} \quad x > x_n \end{cases}$$

dargestellt wird ($F_n'(x) = 0$ für $x \neq x_n$).

6.1.3. Die Ableitung eines Integrals nach der oberen Grenze. Wie wir früher (vgl. 6.1.1.) festgestellt hatten, ist das Integral

$$\int_a^x \varphi(t)\, dt$$

einer beliebigen integrierbaren Funktion als Differenz zweier monotoner Funktionen darstellbar. Daraus ergibt sich mit Satz 1 sofort folgendes Resultat.

Satz 3. *Für jede integrierbare Funktion φ existiert die Ableitung*

$$\frac{d}{dx} \int_a^x \varphi(t)\, dt \tag{14}$$

für fast alle x.

Es muß hier betont werden, daß wir bisher nur die Existenz der Ableitung (14) bewiesen haben. Die Frage, ob die Gleichheit

$$\frac{d}{dx} \int_a^x \varphi(t)\, dt = \varphi(x)$$

besteht, ist bisher nicht betrachtet worden. Wir werden in 6.3. darauf zurückkommen und zeigen, daß diese Gleichung für eine beliebige integrierbare Funktion φ fast überall richtig ist.

6.2. Funktionen von beschränkter Variation

Die Frage nach der Differenzierbarkeit eines Lebesgueschen Integrals nach der oberen Grenze führte uns zur Betrachtung einer Klasse von Funktionen, die als Differenz zweier monotoner Funktionen darstellbar waren. In diesem Abschnitt geben wir für diese Klasse ein andere Beschreibung an, die sich nicht auf den Begriff der Monotonie stützt, und untersuchen grundlegende Eigenschaften dieser Funktionen. Wir beginnen mit zwei notwendigen Definitionen.

Definition 1. Gegeben sei eine auf dem Intervall $[a, b]$ definierte reelle Funktion $f(x)$. Sie heißt *Funktion von beschränkter Variation* (oder *beschränkter Schwankung*), wenn es eine Konstante C gibt, so daß für jede Zerlegung des Intervalls $[a, b]$ durch die Punkte

$$a = x_0 < x_1 < \cdots < x_n = b$$

die Ungleichung

$$\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})| \leqq C \tag{1}$$

erfüllt ist.

Jede monotone Funktion ist von beschränkter Variation, weil für sie die Summe in (1) von der Wahl der Zerlegung unabhängig ist und stets den Wert $|f(b) - f(a)|$ hat.

Definition 2. Ist f eine Funktion von beschränkter Variation auf $[a, b]$, dann heißt die obere Grenze der Summen (1), gebildet für alle möglichen endlichen Zerlegungen des Intervalls $[a, b]$, *vollständige* (*totale*) *Variation* (oder *vollständige Schwankung*) der Funktion f auf $[a, b]$, in Zeichen $V_a^b[f]$, d. h., es ist

$$V_a^b[f] = \sup \sum_{k=1}^{n} |f(x_k) - f(x_{k-1})|.$$

Bemerkung. Ist die Funktion f auf der ganzen Zahlengeraden definiert, so heißt f *Funktion von beschränkter Variation auf* $(-\infty, \infty)$, wenn die Größen

$$V_a^b[f]$$

für alle Intervalle $[a, b]$ als Menge beschränkt sind. Dabei wird dann

$$\lim_{\substack{b \to \infty \\ a \to -\infty}} V_a^b[f]$$

die *vollständige* (oder *totale*) *Variation der Funktion* f auf der Geraden $-\infty < x < \infty$ genannt und mit $V_{-\infty}^{\infty}[f]$ bezeichnet.

Wir beweisen nun einige grundlegende Eigenschaften der vollständigen Variation einer Funktion.

1. *Wenn α eine konstante Zahl ist, dann gilt*

$$V_a^b[\alpha f] = |\alpha| \; V_a^b[f].$$

Das folgt sofort aus der Definition von $V_a^b[f]$.

2. *Wenn f und g Funktionen von beschränkter Variation sind, dann ist $f + g$ auch von beschränkter Variation, und es gilt*

$$V_a^b[f + g] \leqq V_a^b[f] + V_a^b[g]. \tag{2}$$

Für jede Zerlegung des Intervalls $[a, b]$ ist nämlich

$$\sum_k |f(x_k) + g(x_k) - f(x_{k-1}) - g(x_{k-1})|$$
$$\leqq \sum_k |f(x_k) - f(x_{k-1})| + \sum_k |g(x_k) - g(x_{k-1})|,$$

woraus sich wegen

$$\sup (A + B) \leqq \sup A + \sup B$$

die Ungleichung (2) ergibt.

Aus den Eigenschaften 1 und 2 folgt unmittelbar, daß eine Linearkombination aus Funktionen von beschränkter Variation (auf einen festen Intervall $[a, b]$) wieder eine Funktion von beschränkter Variation ist, d. h., *die Funktionen von beschränkter Variation bilden einen linearen Raum* (im Unterschied zur Menge der monotonen Funktionen, die kein linearer Raum ist).

3. *Ist $a < b < c$, dann gilt*

$$V_a^b[f] + V_b^c[f] = V_a^c[f]. \tag{3}$$

Das ergibt sich folgendermaßen. Wir betrachten zuerst eine Zerlegung des Intervalls $[a, c]$, die den Punkt b als Teilpunkt enthält, etwa $x_r = b$. Für diese Zerlegung ist

$$\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})| = \sum_{k=1}^{r} |f(x_k) - f(x_{k-1})| + \sum_{k=r+1}^{n} |f(x_k) - f(x_{k-1})|$$
$$\leqq V_a^b[f] + V_b^c[f]. \tag{4}$$

Ist eine beliebige Zerlegung des Intervalls $[a, c]$ vorgegeben, so kann man zu den Teilpunkten dieser Zerlegung immer noch den Punkt b hinzufügen, ohne daß sich die Summe

$$\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})|$$

dadurch verkleinert. Folglich ist die Ungleichung (4) für jede Zerlegung des Intervalls $[a, c]$ richtig. Daraus folgt sofort

$$V_a^c[f] \leqq V_a^b[f] + V_b^c[f].$$

Andererseits gibt es für jedes $\varepsilon > 0$ Zerlegungen der Intervalle $[a, b]$ und $[b, c]$, so daß

$$\sum_i |f(x_i') - f(x_{i-1}')| > V_a^b[f] - \frac{\varepsilon}{2}$$

und

$$\sum_j |f(x_j'') - f(x_{j-1}'')| > V_b^c[f] - \frac{\varepsilon}{2}$$

ist. Fügen wir diese beiden Zerlegungen zusammen, so ergibt sich eine Zerlegung des Intervalls $[a, c]$, für die

$$\sum_k |f(x_k) - f(x_{k-1})| = \sum_i |f(x_i') - f(x_{i-1}')| + \sum_j |f(x_j'') - f(x_{j-1}'')|$$
$$> V_a^b[f] + V_b^c[f] - \varepsilon$$

gilt. Da $\varepsilon > 0$ beliebig wählbar ist, folgt hieraus

$$V_a^c[f] \geqq V_a^b[f] + V_b^c[f]. \tag{5}$$

Aus (4) und (5) ergibt sich die behauptete Gleichung (3).

Weil die Variation einer beliebigen Funktion auf einem beliebigen Intervall stets nichtnegativ ist, ergibt sich aus Eigenschaft 3 die Eigenschaft

4. *Die Funktion*

$$v(x) = V_a^x[f]$$

ist monoton nichtfallend.

5. *Wenn f in einem Punkt x^* von links stetig ist, dann ist auch $v(x)$ in diesem Punkt von links stetig.*

Auf Grund der Stetigkeit von f gibt es zu jedem vorgegebenen $\varepsilon > 0$ ein $\delta > 0$, so daß $|f(x^*) - f(x)| < \frac{\varepsilon}{2}$ für $x^* - \delta < x \leqq x^*$ ist. Wir wählen nun eine Zerlegung

$$a = x_0 < x_1 < \cdots < x_n = x^*$$

des Intervalls $[x_0, x^*]$ so aus, daß

$$V_a^{x^*}[f] - \sum_{k=1}^{n} |f(x_k) - f(x_{k-1})| < \frac{\varepsilon}{2} \tag{6}$$

ist. Dabei können wir

$$x^* - x_{n-1} < \delta$$

annehmen (andernfalls nehmen wir noch einen Teilpunkt zur Zerlegung hinzu, was die Summe auf der linken Seite von (6) nicht verkleinert). Daraus folgt

$$|f(x^*) - f(x_{n-1})| < \frac{\varepsilon}{2}$$

und

$$V_a^{x^*}[f] - \sum_{k=1}^{n-1} |f(x_k) - f(x_{k-1})| < \varepsilon.$$

Damit ist aber erst recht

$$V_a^{x^*}[f] - V_a^{x_{n-1}}[f] < \varepsilon,$$

d. h.

$$v(x^*) - v(x_{n-1}) < \varepsilon.$$

Da v eine monoton nichtfallende Funktion ist, ergibt sich aus dieser Ungleichung, daß $v(x^*) - v(x) < \varepsilon$ für alle $x \in [x_{n-1}, x^*]$ ist, was zu zeigen war.

Wenn f in einem Punkt x^* von rechts stetig ist, dann zeigen analoge Überlegungen, daß auch v in diesem Punkt von rechts stetig ist. Damit gilt: *Wenn f in einem Punkt (oder auf dem ganzen Intervall $[a, b]$) stetig ist, dann ist dort auch v stetig.*

Ist f eine beliebige Funktion von beschränkter Variation auf $[a, b]$ und v ihre vollständige Variation auf $[a, x]$, dann ist die Differenz

$$\varphi = v - f$$

eine monoton nichtfallende Funktion. Denn für $x' \leqq x''$ ist stets

$$|f(x'') - f(x')| \leqq v(x'') - v(x') = V_{x'}^{x''}[f],$$

und daher ist die Differenz

$$\varphi(x'') - \varphi(x') = [v(x'') - v(x')] - [f(x'') - f(x')] \tag{7}$$

stets positiv. Da andererseits die Funktion f in der Form

$$f = v - \varphi$$

darstellbar ist, erhalten wir folgendes Resultat.

Satz 1. *Jede Funktion von beschränkter Variation kann als Differenz zweier monoton nichtfallender Funktionen dargestellt werden.*

Die Umkehrung dieses Satzes ist offensichtlich auch richtig: Jede Funktion, die als Differenz zweier monotoner Funktionen dargestellt werden kann, ist von beschränkter Variation. Daher stimmt die Gesamtheit der Funktionen, die als Differenz monotoner Funktionen darstellbar sind, mit der Gesamtheit der Funktionen von beschränkter Variation überein.

Aus Satz 1 und dem in 6.1.2. bewiesenen Satz von LEBESGUE über die Differenzierbarkeit monotoner Funktionen folgt nun unmittelbar:

Jede Funktion von beschränkter Variation besitzt fast überall eine endliche Ableitung.

Aufgaben

1. Man zeige: Wenn f auf $[a, b]$ eine beschränkte Ableitung besitzt (d. h., $f'(x)$ existiert überall, und es ist $|f'(x)| < C$, $x \in [a, b]$), dann ist f eine Funktion von beschränkter Variation, und es gilt

$$V_a^b[f] \leqq C(b - a).$$

2. Man zeige, daß die vollständige Variation der Funktion $f(x) = x \sin \frac{1}{x}$ auf $[0, 1]$ unendlich ist.

Man überlegt sich leicht, daß konstante Funktionen und nur diese eine verschwindende vollständige Variation besitzen. Setzen wir

$$\|f\| = V_a{}^b[f],$$

dann besitzt diese Größe die Eigenschaften 2 und 3 einer Norm (vgl. 3.3.1.), während die Eigenschaft 1 entsprechend der vorhergehenden Bemerkung nicht erfüllt ist. Werden jedoch nur Funktionen auf $[a, b]$ betrachtet, die der Bedingung $f(a) = 0$ genügen, dann besitzt die Größe $V_a{}^b[f]$ alle Eigenschaften einer Norm, und es ist folgende Definition sinnvoll. Die Menge $V^0[a, b]$ der Funktionen von beschränkter Variation auf dem Intervall $[a, b]$, die der Bedingung $f(a) = 0$ genügen, bildet mit der wie üblich definierten Addition von Funktionen und Multiplikation mit reellen Zahlen und der Norm

$$\|f\| = V_a{}^b[f]$$

einen linearen normierten Raum, der als *Raum der Funktionen von beschränkter Variation* bezeichnet wird. (Der Leser beweise die Vollständigkeit dieses Raumes!)

6.3. Die Ableitung des unbestimmten Lebesgueschen Integrals

In 6.1. haben wir bewiesen, daß das Lebesguesche Integral

$$\int_a^x f(t)\,dt$$

als Funktion von x fast überall eine endliche Ableitung besitzt, jedoch nicht gezeigt, wie diese Ableitung mit dem Integranden zusammenhängt. Jetzt beweisen wir folgendes Resultat, auf das wir schon am Ende von 6.1. hingewiesen haben.

Satz 1. *Für jede integrierbare Funktion f gilt fast überall*

$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x).$$

Beweis. Wir setzen

$$\Phi(x) = \int_a^x f(t)\,dt$$

und zeigen zuerst, daß für fast alle x

$$f(x) \geqq \Phi'(x)$$

gilt. Ist $f(x) < \Phi'(x)$, dann gibt es zwei rationale Zahlen α und β, so daß

$$f(x) < \alpha < \beta < \Phi'(x) \tag{1}$$

gilt. Bezeichnen wir mit $E_{\alpha\beta}$ die Menge aller Punkte x, für die (1) erfüllt ist, dann ist $E_{\alpha\beta}$ meßbar, da f und Φ' meßbar sind. Im weiteren zeigen wir, daß $\mu(E_{\alpha\beta}) = 0$ für beliebige rationale Zahlen α, β ist. Auf Grund von

$$\{x: f(x) < \Phi'(x)\} \subset \bigcup_{\alpha,\beta} E_{\alpha\beta}$$

und der Abzählbarkeit der rationalen Zahlen folgt daraus

$$\mu\{x: f(x) < \Phi'(x)\} = 0,$$

was mit der obigen Behauptung äquivalent ist.

Es sei nun eine beliebige Zahl $\varepsilon > 0$ vorgegeben und $\delta > 0$ so bestimmt, daß

$$\left|\int_e f(t)\,dt\right| < \varepsilon$$

für $\mu(e) < \delta$ ist (ein solches δ existiert auf Grund der absoluten Stetigkeit des Integrals). G sei eine offene Menge in $[a, b]$ mit der Eigenschaft

$$G \supset E_{\alpha\beta} \quad \text{und} \quad \mu(G) < \mu(E_{\alpha\beta}) + \delta.$$

Ist nun $x \in E_{\alpha\beta}$, dann gilt für alle genügend nahe bei x gelegenen $\xi > x$

$$\frac{\Phi(\xi) - \Phi(x)}{\xi - x} > \beta \tag{2}$$

oder

$$\Phi(\xi) - \beta\xi > \Phi(x) - \beta x,$$

d. h., x ist ein von rechts unsichtbarer Punkt für die Funktion $\Phi(x) - \beta x$ auf einem beliebigen Intervall aus der Darstellung von G durch paarweise disjunkte offene Intervalle. Nach dem Lemma von F. Riesz (vgl. 6.1.2.) gibt es nun eine offene Menge $S = \bigcup_k (a_k, b_k)$ mit den Eigenschaften $E_{\alpha\beta} \subset S \subset G$ und

$$\Phi(b_k) - \beta b_k \geqq \Phi(a_k) - \beta a_k.$$

Nach dieser Ungleichung ist

$$\Phi(b_k) - \Phi(a_k) \geqq \beta(b_k - a_k),$$

d. h.

$$\int_{a_k}^{b_k} f(t)\,dt \geqq \beta(b_k - a_k).$$

Summieren wir diese Ungleichungen für alle Intervalle der Darstellung von S, dann erhalten wir

$$\int_S f(t)\,dt \geqq \beta\mu(S). \tag{3}$$

Andererseits ist

$$\int_S f(t)\,dt = \int_{E_{\alpha\beta}} f(t)\,dt + \int_{S\setminus E_{\alpha\beta}} f(t)\,dt < \alpha\mu(E_{\alpha\beta}) + \varepsilon \leqq \alpha\mu(S) + \varepsilon + |\alpha|\,\delta. \tag{4}$$

Ein Vergleich von (3) und (4) ergibt

$$\beta\mu(S) \leqq \alpha\mu(S) + \varepsilon + |\alpha|\delta,$$

d. h.

$$\mu(S) \leqq \frac{\varepsilon + |\alpha|\delta}{\beta - \alpha}.$$

Daraus folgt: Die Menge $E_{\alpha\beta}$ kann durch eine offene Menge S von beliebig kleinem Maß überdeckt werden (δ kann z. B. so bestimmt werden, daß $|\alpha|\delta \leqq \varepsilon$ ist), d. h. $\mu(E_{\alpha\beta}) = 0$.

Nach der Bemerkung am Anfang des Beweises ist bisher gezeigt, daß fast überall

$$f(x) \geqq \Phi'(x)$$

gilt. Ersetzen wir $f(x)$ durch $-f(x)$, so erhalten wir auf dieselbe Weise, daß fast überall

$$-f(x) \geqq -\Phi'(x),$$

d. h.

$$f(x) \leqq \Phi'(x)$$

ist. Aus beiden Ungleichungen folgt

$$f(x) = \Phi'(x) = \frac{d}{dx}\int_a^x f(t)\,dt$$

für fast alle $x \in [a, b]$, was zu zeigen war.

6.4. Berechnung einer Funktion aus ihrer Ableitung. Absolut stetige Funktionen

Bisher haben wir die erste der am Anfang dieses Kapitels gestellten Fragen gelöst; wir haben gezeigt, daß die Gleichheit

$$\frac{d}{dx}\int_x^a f(t)\,dt = f(x) \qquad \text{(fast überall)}$$

für eine beliebige auf $[a, b]$ integrierbare Funktion besteht. Nun betrachten wir die zweite der dort formulierten Fragen, d. h., wir untersuchen, ob sich die im Fall

stetig differenzierbarer Funktionen gut bekannte Newton-Leibnizsche Formel

$$F(x) = F(a) + \int_a^x F'(t)\, dt \tag{1}$$

auf den allgemeinen Fall des Lebesgueschen Integrals ausdehnen läßt. Dabei ist es natürlich, die Betrachtungen auf fast überall differenzierbare Funktionen einzuschränken, da sonst die Gleichung (1) keinen Sinn hat. Wie wir bereits wissen, sind solche Funktionen insbesondere die Funktionen von beschränkter Variation. Andererseits ist auch das Integral auf der rechten Seite von (1) eine Funktion von beschränkter Variation. Daher kann die Gleichung (1) nur in der Klasse der Funktionen von beschränkter Variation und nicht auf einer noch umfangreicheren Klasse richtig sein. Weil jede Funktion von beschränkter Variation als Differenz zweier monoton nichtfallender Funktionen darstellbar ist, werden wir in erster Linie monotone Funktionen betrachten.

Satz 1. *Die Ableitung f' einer monoton nichtfallenden Funktion f ist integrierbar, und es gilt*

$$\int_a^x f'(x)\, dx \leqq f(b) - f(a).$$

Beweis. Die Ableitung der Funktion f im Punkt x ist als Grenzwert des Quotienten[1])

$$\varphi_h(x) = \frac{f(x+h) - f(x)}{h} \tag{2}$$

für $h \to 0$ definiert. Aus der Monotonie der Funktion f folgt ihre Integrierbarkeit und die Integrierbarkeit jeder der Funktionen φ_h. Daher kann die Gleichung (2) integriert werden. Man erhält

$$\int_a^b \varphi_h(x)\, dx = \frac{1}{h}\int_a^b f(x+h)\, dx - \frac{1}{h}\int_a^b f(x)\, dx$$
$$= \frac{1}{h}\int_b^{b+h} f(x)\, dx - \frac{1}{h}\int_a^{a+h} f(x)\, dx.$$

Der Ausdruck auf der rechten Seite dieser Gleichung strebt für $h \to 0 + 0$ gegen $f(b) - f(a + 0)$. Daher erhalten wir nach dem Satz von Fatou

$$\int_a^b f'(x)\, dx \leqq \lim_{h\to 0} \int_a^b \varphi_h(x)\, dx = f(b) - f(a+0) \leqq f(b) - f(a),$$

[1]) Damit der Ausdruck $f(x + h)$ für jeden Punkt $x \in [a, b]$ sinnvoll ist, nehmen wir $f(x) = f(b)$ für $x > b$ und $f(x) = f(a)$ für $x < a$ an.

wobei auch die Integrierbarkeit von f' aus diesem Satz folgt. Damit ist Satz 1 bewiesen.

Es ist nicht schwierig, eine monotone Funktion f zu finden, für die echte Ungleichheit

$$\int_a^b f'(x)\,dx < f(b) - f(a)$$

besteht. Die Funktion

$$f(x) = \begin{cases} 0 \text{ für } 0 \leqq x \leqq \frac{1}{2}, \\ 1 \text{ für } \frac{1}{2} < x \leqq 1 \end{cases}$$

ist z. B. eine solche Funktion. Interessant ist, daß es sogar stetige monotone Funktionen gibt, für welche die echte Ungleichheit

$$\int_a^x f'(t)\,dt < f(x) - f(a)$$

bei beliebigem x besteht. Eines der einfachsten Beispiele hierfür ist das folgende. Wir betrachten auf dem Intervall $[0, 1]$ die Cantorsche Menge und definieren die Funktion f auf den Zwischenintervallen n-ten Ranges (einschließlich ihrer Grenzen) durch

$$f(t) = \frac{2k-1}{2^n}, \quad k = 1, 2, 3, \ldots, 2^{n-1}, \tag{3}$$

wobei diese Intervalle von links nach rechts zu numerieren sind. Es ist also

$$f(t) = \frac{1}{2} \text{ für } \frac{1}{3} \leqq t \leqq \frac{2}{3},$$

$$f(t) = \frac{1}{4} \text{ für } \frac{1}{9} \leqq t \leqq \frac{2}{9},$$

$$f(t) = \frac{3}{4} \text{ für } \frac{7}{9} \leqq t \leqq \frac{8}{9},$$

usw. (vgl. Abb. 21). Damit ist f auf dem gesamten Intervall $[0, 1]$ mit Ausnahme der Punkte zweiter Art der Cantorschen Menge (das sind die Punkte aus $[0, 1]$, die weder zu den Zwischenintervallen gehören noch Randpunkte von solchen sind) definiert. In den letztgenannten Punkten definieren wir f auf folgende Weise. Ist t^* ein Punkt zweiter Art der Cantorschen Menge, $\{t_m\}$ eine wachsende und $\{t_m'\}$ eine

fallende gegen t^* konvergente Folge von Punkten erster Art (Endpunkte von Zwischenintervallen), dann existieren die Grenzwerte

$$\lim_{m\to\infty} f(t_m) \quad \text{und} \quad \lim_{m\to\infty} f(t_m') \tag{4}$$

und sind, wie man leicht nachprüft, gleich. Ihren gemeinsamen Wert setzen wir als Funktionswert $f(t^*)$ im Punkt t^* fest. Die so definierte Funktion wird als *Cantorsche Treppenfunktion* bezeichnet. Sie ist auf dem ganzen Intervall $[0, 1]$ stetig und mono-

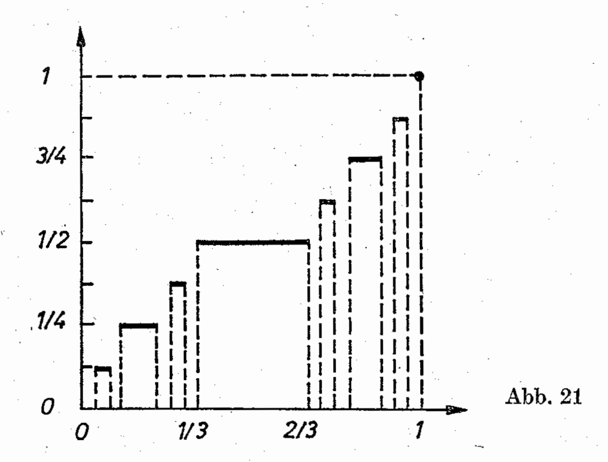


Abb. 21

ton nichtfallend. Ihre Ableitung ist in jedem Punkt eines Zwischenintervalls offensichtlich gleich 0 und daher auf $[0, 1]$ fast überall 0, so daß in diesem Fall

$$0 = \int_a^x f'(t)\,dt < f(x) - f(0) = f(x)$$

für ein beliebiges x aus $(0, 1]$ gilt.

Wir weisen nebenbei darauf hin, daß für eine monotone Funktion $f(x)$ aus der Gleichheit $\int_a^b f(t)\,dt = f(b) - f(a)$ die Gleichheit $\int_a^x f'(t)\,dt = f(x) - f(a)$ für jedes $x \in (a, b]$ folgt.

Um die Klasse der Funktionen beschreiben zu können, für die die Gleichheit

$$\int_a^b f'(t)\,dt = f(b) - f(a)$$

besteht, führen wir folgende Definition ein.

Definition 1. Eine auf dem Intervall $[a, b]$ definierte Funktion heißt dort *absolut stetig*, wenn es für ein beliebiges $\varepsilon > 0$ stets ein $\delta > 0$ gibt, so daß für jedes endliche

System paarweise disjunkter Intervalle

$$(a_k, b_k), \quad k = 1, 2, \ldots, n,$$

mit der Eigenschaft

$$\sum_{k=1}^{n} (b_k - a_k) < \delta$$

die Ungleichung

$$\sum_{k=1}^{n} |f(b_k) - f(a_k)| < \varepsilon$$

erfüllt ist.

Es ist klar, daß jede absolut stetige Funktion gleichmäßig stetig ist. Die Umkehrung dieser Aussage ist jedoch im allgemeinen nicht richtig. Zum Beispiel ist die Fortsetzung der Cantorschen Treppenfunktion, die links von 0 identisch 0 und rechts von 1 identisch 1 ist, stetig auf dem Intervall $[-1, 2]$ und demzufolge gleichmäßig stetig, jedoch nicht absolut stetig. Das ergibt sich daraus, daß die Cantorsche Menge durch ein endliches System von Intervallen (a_k, b_k), $k = 1, 2, \ldots, n$, mit beliebig kleiner Gesamtlänge überdeckt werden kann, während offensichtlich

$$\sum_{k=1}^{n} |f(b_k) - f(a_k)| = 1$$

ist.

Wir beweisen nun einige grundlegende Eigenschaften absolut stetiger Funktionen.

1. In Definition 1 kann an Stelle eines beliebigen endlichen Systems von Intervallen mit einer Gesamtlänge kleiner als δ ein beliebiges abzählbares System von Intervallen mit einer Gesamtlänge kleiner als δ betrachtet werden, ohne daß sich die Definition inhaltlich ändert. Denn kann man für ein vorgegebenes $\varepsilon > 0$ ein $\delta > 0$ angeben, so daß

$$\sum_{k=1}^{n} |f(b_k) - f(a_k)| < \varepsilon$$

für ein beliebiges endliches System von Intervallen (a_k, b_k) mit

$$\sum_{k=1}^{n} (b_k - a_k) < \delta$$

gilt, dann gilt auch für ein beliebiges abzählbares System von Intervallen (α_k, β_k), $k = 1, 2, \ldots,$ mit

$$\sum_{k=1}^{\infty} (\beta_k - \alpha_k) < \delta$$

für beliebiges n stets

$$\sum_{k=1}^{n} |f(\beta_k) - f(\alpha_k)| < \varepsilon$$

und, wenn man den Grenzübergang $n \to \infty$ durchführt,

$$\sum_{k=1}^{\infty} |f(\beta_k) - f(\alpha_k)| \leqq \varepsilon.$$

2. *Jede absolut stetige Funktion ist von beschränkter Variation.*

Absolute Stetigkeit auf dem Intervall $[a, b]$ bedeutet für die Funktion f insbesondere, daß für jedes $\varepsilon > 0$ ein $\delta > 0$ gefunden werden kann, so daß die vollständige Variation der Funktion f auf einem Intervall mit einer Länge kleiner als δ nicht größer als ε ist. Da das Intervall $[a, b]$ in endlich viele Intervalle mit einer Länge kleiner als δ zerlegt werden kann, ist dann auch die vollständige Variation von f auf $[a, b]$ endlich.

3. *Die Summe von absolut stetigen Funktionen und das Produkt einer solchen Funktion mit einer Zahl sind wieder absolut stetige Funktionen.*

Das folgt sofort aus der Definition der absoluten Stetigkeit und den Eigenschaften des absoluten Betrages einer Summe bzw. eines Produktes.

Die Eigenschaften 2 und 3 besagen, daß *die absolut stetigen Funktionen im Raum aller Funktionen von beschränkter Variation eine lineare Menge bilden.*

4. *Jede absolut stetige Funktion kann als Differenz zweier absolut stetiger nichtfallender Funktionen dargestellt werden.*

Da jede absolut stetige Funktion f als Differenz

$$f = v - g$$

der beiden nichtfallenden Funktionen

$$v(x) = V_a^x[f] \quad \text{und} \quad g(x) = v(x) - f(x)$$

dargestellt werden kann, genügt es, für den Nachweis der Eigenschaft 4 zu zeigen, daß die Funktion $v(x)$ absolut stetig ist. Die absolute Stetigkeit von g folgt dann sofort aus Eigenschaft 3. Es sei ein $\varepsilon > 0$ vorgegeben und dazu ein $\delta > 0$ entsprechend der Definition der absoluten Stetigkeit von f bestimmt. Für ein beliebiges System von n Intervallen (a_k, b_k) mit einer Gesamtlänge kleiner als δ betrachten wir die Summe

$$\sum_{k=1}^{n} \big(v(b_k) - v(a_k)\big). \tag{5}$$

Die Summe ist gleich der oberen Grenze der Zahlen

$$\sum_{k=1}^{n} \sum_{l=1}^{m_k} |f(x_{k,l}) - f(x_{k,l-1})| \tag{6}$$

bei allen möglichen endlichen Zerlegungen

$$a_1 = x_{1,0} < x_{1,1} < x_{1,2} < \cdots < x_{1,m_1} = b_1;$$
$$a_2 = x_{2,0} < x_{2,1} < x_{2,2} < \cdots < x_{2,m_2} = b_2,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$a_n = x_{n,0} < x_{n,1} < x_{n,2} < \cdots < x_{n,m_n} = b_n$$

der Intervalle $(a_1, b_1), \ldots, (a_n, b_n)$. Da die Gesamtlänge aller Intervalle $(x_{k,l-1}, x_{k,l})$ für jede Zerlegung kleiner als δ ist, wird auf Grund der absoluten Stetigkeit von f jede der Summen (6) kleiner als ε sein. Folglich ist auch ihre obere Grenze (5) nicht größer als ε, was die absolute Stetigkeit der Funktion v bedeutet.

Die beiden folgenden Sätze beschreiben den engen Zusammenhang zwischen dem Begriff der absoluten Stetigkeit und dem unbestimmten Lebesgueschen Integral.

Satz 2. *Das unbestimmte Integral*

$$F(x) = \int_a^x f(t)\,dt$$

einer integrierbaren Funktion f ist eine absolut stetige Funktion seiner oberen Grenze.

Beweis. Ist $\{(a_k, b_k)\}$ ein beliebiges endliches System paarweise disjunkter Intervalle, so ist

$$\sum_{k=1}^{n} |F(b_k) - F(a_k)| = \sum_{k=1}^{n} \left| \int_{a_k}^{b_k} f(t)\,dt \right| \leqq \sum_{k=1}^{n} \int_{a_k}^{b_k} |f(t)|\,dt = \int_{\bigcup_k (a_k,b_k)} |f(t)|\,dt.$$

Strebt die Gesamtlänge des Intervallsystems gegen 0, dann folgt aus der absoluten Stetigkeit des Lebesgueschen Integrals (vgl. Satz 5 aus 5.5.), daß auch das Integral gegen 0 strebt.

Satz 3 (Lebesgue). *Die Ableitung $f = F'$ einer auf dem Intervall $[a, b]$ absolut stetigen Funktion F ist integrierbar, und es gilt für jedes $x \in [a, b]$*

$$\int_a^x f(t)\,dt = F(x) - F(a).$$

Die Sätze 2 und 3 zeigen, daß sich die absolut stetigen Funktionen und nur diese aus ihrer Ableitung durch Integration eindeutig (bis auf eine additive Konstante) berechnen lassen.

Für den Beweis des Satzes 3 brauchen wir folgendes Lemma.

Lemma. *Wenn die Ableitung einer absolut stetigen monoton nichtfallenden Funktion f fast überall gleich 0 ist, dann ist diese Funktion eine Konstante.*

Beweis des Lemmas. Weil f eine stetige monoton nichtfallenden Funktion auf $[a, b]$ ist, stimmt ihr Wertevorrat mit dem Intervall $[f(a), f(b)]$ überein. Daher genügt

es zu zeigen, daß dieses Intervall die Länge 0 hat, wenn fast überall $f'(x) = 0$ ist. Es sei $E = \{x : x \in [a, b],\ f'(x) = 0\}$, $Z = [a, b] \setminus E$. Dann ist nach Voraussetzung $\mu(Z) = 0$. Für ein beliebig vorgegebenes $\varepsilon > 0$ bestimmen wir nun ein $\delta > 0$ entsprechend der Definition der absoluten Stetigkeit von f und wählen eine Überdeckung der Menge Z durch ein System offener Intervalle (a_k, b_k) mit einer Gesamtlänge kleiner als δ. Das ist wegen $\mu(Z) = 0$ möglich. Entsprechend der Auswahl von δ ist

$$\sum_k |f(b_k) - f(a_k)| < \varepsilon,$$

d. h., die Menge der Funktionswerte von f auf dem System $\{(a_k, b_k)\}$ (und damit erst recht auf der Untermenge Z) hat ein Maß kleiner als ε. Da $\varepsilon > 0$ beliebig war, folgt $\mu\big(f(Z)\big) = 0$.

Wir betrachten nun die Menge $E = [a, b] \setminus Z$. Ist $x_0 \in E$, so ist $f'(x_0) = 0$, und für alle hinreichend nahe bei x_0 liegenden x ist die Ungleichung

$$\frac{f(x) - f(x_0)}{x - x_0} < \varepsilon$$

erfüllt. Nehmen wir $x > x_0$ an, so heißt das

$$f(x) - f(x_0) < \varepsilon(x - x_0)$$

oder

$$\varepsilon x_0 - f(x_0) < \varepsilon x - f(x),$$

d. h., x_0 ist ein von rechts unsichtbarer Punkt für die Funktion $g(x) = \varepsilon x - f(x)$. Nach dem Lemma von F. RIESZ ist die Menge E damit in einem endlichen oder abzählbaren System von Intervallen (α_k, β_k) enthalten, in deren Endpunkten die Bedingungen

$$\varepsilon\beta_k - f(\beta_k) \geqq \varepsilon\alpha_k - f(\alpha_k),$$

also

$$f(\beta_k) - f(\alpha_k) \leqq \varepsilon(\beta_k - \alpha_k)$$

erfüllt sind. Summieren wir die letzten Ungleichungen über k, so ergibt sich

$$\sum_k \big(f(\beta_k) - f(\alpha_k)\big) \leqq \varepsilon \sum_k (\beta_k - \alpha_k) \leqq \varepsilon(b - a).$$

Das bedeutet, daß die Menge der Funktionswerte von f auf einem E überdeckenden Intervallsystem ein Maß kleiner als $\varepsilon(b - a)$ hat. Da $\varepsilon > 0$ beliebig wählbar ist, folgt $\mu\big(f(E)\big) = 0$.

Damit ist gezeigt, daß $f(E)$ und $f(Z)$ das Maß 0 haben. Da aber die Vereinigung beider Mengen gleich $[f(a), f(b)]$ ist, hat auch dieses Intervall das Maß 0, was zu zeigen war.

Auf der Gundlage des Lemmas ist nun Satz 3 leicht zu beweisen. Wir beschränken uns dabei auf den Fall, daß $F(x)$ nichtfallend ist. In diesem Fall ist

$$\Phi(x) = F(x) - \int_a^x f(t)\,dt$$

auch eine monoton nichtfallende Funktion, denn für $x'' > x'$ ist

$$\Phi(x'') - \Phi(x') = F(x'') - F(x') - \int_{x'}^{x''} f(t)\,dt \geqq 0.$$

Außerdem ist Φ absolut stetig (als Differenz zweier absolut stetiger Funktionen) und $\Phi'(x) = 0$ fast überall (nach Satz 1 aus 6.3.). Daher ist Φ nach dem obigen Lemma eine Konstante. Setzen wir in (7) $x = a$, so sehen wir, daß diese Konstante gleich $F(a)$ ist. Damit ist der Satz bewiesen.

Wir haben früher schon festgestellt, daß jede Funktion f von beschränkter Variation als Summe

$$f = H + \varphi$$

aus einer Sprungfunktion H und einer stetigen Funktion φ von beschränkter Variation dargestellt werden kann. Setzt man für die stetige (im allgemeinen nicht absolut stetige) Funktion φ

$$\psi(x) = \int_a^x \varphi'(t)\,dt,$$

so ist die Differenz

$$\chi = \varphi - \psi$$

ebenfalls eine Funktion von beschränkter Variation mit fast überall verschwindender Ableitung

$$\frac{d}{dx}\,\chi(x) = \varphi'(x) - \frac{d}{dx}\int_a^x \varphi'(t)\,dt = 0.$$

Bezeichnen wir eine stetige Funktion von beschränkter Variation mit fast überall verschwindender Ableitung als *singuläre* Funktion, so können wir folgendes Resultat formulieren:

Jede Funktion von beschränkter Variation kann in Form einer Summe

$$f = H + \psi + \chi \tag{8}$$

aus einer Sprungfunktion, einer absolut stetigen Funktion und einer singulären Funktion dargestellt werden.

Es läßt sich leicht zeigen, daß jeder der Summanden in der Zerlegung (8) durch die Funktion f bis auf eine additive Konstante eindeutig bestimmt ist. Werden alle Funktionen in (8) durch die Forderung normiert, daß sie in dem Punkt $x = a$ verschwinden sollen, dann ist die Zerlegung (8) eindeutig. Differenziert man die Gleichung (8), so ergibt sich

$$f'(x) = \varphi'(x)$$

für fast alle x, da H' und χ' fast überall gleich 0 sind. Daher kann bei der Integration der Ableitung einer Funktion von beschränkter Variation diese Funktion durch ihre absolut stetige Komponente ersetzt werden. Die zwei anderen Komponenten (die Sprungfunktion und die singuläre Funktion) „verschwinden dabei spurlos".

Es ist sehr interessant, die Resultate dieses Abschnitts mit denen der Theorie der Distributionen zu vergleichen. Wie in Kapitel 4 werden wir unter einer Distribution ein lineares stetiges Funktional auf dem Raum K der finiten unendlichen oft differenzierbaren Funktionen verstehen. Dabei entspricht einer gewöhnlichen lokal integrierbaren Funktion f das Funktional, das jedem Element $\varphi \in K$ die Zahl

$$(f, \varphi) = \int_{-\infty}^{\infty} f(x)\, \varphi(x)\, dx$$

zuordnet. Die Distributionsableitung dieses Funktionals ist dann das Funktional, das jedem Element $\varphi \in K$ die Zahl

$$(f', \varphi) = -\int_{-\infty}^{\infty} f(x)\, \varphi'(x)\, dx$$

zuordnet. Da in der Klasse der Distributionen die Gleichung $y' = 0$ nur die üblichen Lösungen (Konstanten) besitzt, ist jede Distribution durch ihre Ableitung bis auf eine additive Konstante bestimmt. Insbesondere ist also jede lokal integrierbare Funktion f (bis auf eine additive Konstante) fast überall durch ihre Distributionsableitung f' bestimmt. Setzen wir voraus, daß die Funktion f fast überall eine lokal integrierbare Ableitung besitzt (das ist erfüllt, wenn f z. B. eine monotone Funktion ist) und bezeichnen diese (übliche) Ableitung mit $f_1 = df/dx$, so können wir diese Funktion als Distribution

$$(f_1, \varphi) = \int_{-\infty}^{\infty} \frac{df}{dx}\, \varphi(x)\, dx$$

auffassen. Man sieht nun leicht, daß diese Distribution f_1 im allgemeinen nicht mit der Distributionsableitung f' übereinstimmt. So ist z. B. für

$$f(x) = \begin{cases} 1 \text{ für } x > 0, \\ 0 \text{ für } x \leqq 0 \end{cases}$$

$f_1 = 0$, aber $f' = \delta$ (vgl. Beispiel 1 in 4.4.3.).

Satz 3 klärt die Beziehung zwischen Distributionsableitung und gewöhnlicher Ableitung einer Funktion. Nach diesem Satz stimmt in der Klasse der Funktionen von beschränkter Variation für absolut stetige Funktionen (und nur für sie!) die gewöhnliche Ableitung mit der Distributionsableitung überein (vgl. die folgende Aufgabe 3).

Die obigen Betrachtungen zeigen uns eine ähnliche Situation, wie wir sie schon in 4.4. besprochen haben: Um grundlegende Operationen der Analysis (hier die Darstellung einer Funktion durch ihre Ableitung) zu sichern, muß man entweder innerhalb des Rahmens der klassischen Analysis die betrachtete Funktionenklasse genügend einschränken (hier absolut stetige Funktionen) oder umgekehrt den Begriff der Funktion (und entsprechend der Ableitung) wesentlich erweitern.

Aufgaben

1. Man zeige, daß die Definition der absoluten Stetigkeit einer Funktion mit der folgenden äquivalent ist: f heißt absolut stetig auf $[a, b]$, wenn jede Untermenge einer Menge vom Maß 0 durch f in eine Menge übergeführt wird, die wieder das Maß 0 hat.

2. Man bestimme die Distributionsableitung der Cantorschen Treppenfunktion.

3. Es sei f eine Funktion von beschränkter Variation, f' ihre Distributionsableitung, f_1 die Distribution, die der gewöhnlichen Ableitung df/dx der Funktion f entspricht. Man zeige, daß

a) $f' = f_1$ gilt, wenn f absolut stetig ist;

b) $f(x)$ fast überall mit einer absolut stetigen Funktion übereinstimmt, wenn $f' = f_1$ gilt, und insbesondere $f(x)$ absolut stetig ist, wenn $f' = f_1$ und f stetig ist.

6.5. Das Lebesguesche Integral als Mengenfunktion. Der Satz von Radon-Nikodym

6.5.1. Reellwertige Maße. Hahnsche und Jordansche Zerlegung eines Maßes. Die in den vergangenen Abschnitten für Funktionen auf der Zahlengeraden dargestellten Begriffe und Aussagen können weitgehend auf Funktionen ausgedehnt werden, die auf einer beliebigen Menge mit Maß gegeben sind. Im folgenden wird diese Ausdehnung dargestellt.

Es sei X eine Menge mit dem (endlichen) Maß μ und f eine auf X bezüglich μ integrierbare Funktion. Dann ist f auch auf jeder meßbaren Untermenge A von X integrierbar und folglich das Integral

$$\Phi(A) = \int_A f(x)\, d\mu \tag{1}$$

(bei festem f) eine Mengenfunktion auf der σ-Algebra $\mathfrak{S}_\mu$ aller meßbaren Untermengen der Menge X. Für eine beliebige Zerlegung

$$A = \bigcup_k A_k$$

der meßbaren Menge A in endlich oder abzählbar viele paarweise disjunkte Mengen A_k gilt (vgl. 5.5.4.)

$$\Phi(A) = \sum_k \Phi(A_k),$$

d. h., die durch (1) definierte Funktion Φ besitzt alle Eigenschaften eines σ-additiven Maßes mit eventueller Ausnahme der Nichtnegativität (wenn f nichtnegativ ist, ist auch Φ nichtnegativ).

Definition 1. Eine beliebige (endliche) σ-additive Mengenfunktion Φ auf einer σ-Algebra von Untermengen einer gegebenen Menge X heißt *reellwertiges Maß*[1]).

Der Begriff des reellwertigen Maßes ist eine natürliche Verallgemeinerung des Begriffes des üblichen (nichtnegativen) σ-additiven Maßes und reduziert sich, wie wir unten sehen werden, in gewissem Sinn auf den üblichen Maßbegriff.

Aufgabe. Man zeige, daß für ein beliebiges (endliches) reellwertiges Maß Φ auf einer σ-Algebra $\mathfrak{S}$ von Mengen eine Konstante c existiert, so daß $|\Phi(A)| \leqq c$ für alle $A \in \mathfrak{S}$ gilt.

Wir führen nun folgende Begriffe ein. Für ein reellwertiges Maß Φ, das auf einer σ-Algebra $\mathfrak{S}$ von Untermengen der Menge X definiert ist, heißt eine Menge $E \in \mathfrak{S}$ *negativ* bezüglich Φ, wenn $\Phi(E \cap F) \leqq 0$ für alle $F \in \mathfrak{S}$ ist; analog heißt E *positiv*, wenn $\Phi(E \cap F) \geqq 0$ für alle $F \in \mathfrak{S}$ ist.

Satz 1. *Ist Φ ein reellwertiges Maß auf X,*[2]) *dann existiert eine solche meßbare Menge $A^- \subset X$, so daß A^- negativ und $A^+ = X \setminus A^-$ positiv (bezüglich Φ) ist.*

Beweis. Wir setzen

$$a = \inf \Phi(A),$$

wobei das Infimum über alle negativen Mengen A gebildet wird. Ist $\{A_n\}$ eine Folge negativer Mengen mit

$$\lim_{n\to\infty} \Phi(A_n) = a,$$

dann ist $A^- = \bigcup_n A_n$, wobei $\Phi(A^-) = a$ ist, die gesuchte Menge. Dafür genügt es zu zeigen, daß $A^+ = X \setminus A^-$ positiv ist. Nehmen wir an, A^+ sei nicht positiv, dann enthält A^+ eine meßbare Untermenge C_0 mit $\Phi(C_0) < 0$. Dabei kann die Menge C_0 selbst nicht negativ sein, weil wir sonst nach Vereinigung mit A^- eine negative Menge $\tilde{A}$ erhielten, für die

$$\Phi(\tilde{A}) < a$$

im Widerspruch zur Konstruktion von A^- wäre. Daher existiert eine kleinste ganze Zahl $k_1 > 0$, für die es eine Untermenge C_1 von C_0 gibt, die die Bedingung

$$\Phi(C_1) \geqq \frac{1}{k_1}$$

---

[1]) Anm. d. Übers.: Im Original wird ein reellwertiges Maß als „Ladung" (заряд) bezeichnet. Damit wird auf die Analogie zu realen elektrischen Ladungen hingewiesen. Eine derartige elektrische Ladung kann man in einen positiven und in einen negativen Anteil zerlegen. Das Analogon hierzu für reellwertige Maße wird in Satz 1 bewiesen.

[2]) Anm. d. Übers.: Die Aussage „Φ ist ein reellwertiges Maß *auf* X" bedeutet (vgl. Definition 1), daß in X eine σ-Algebra $\mathfrak{S}$ von Untermengen ausgezeichnet ist, auf der die Funktion Φ erklärt ist. Eine Menge $A \in \mathfrak{S}$ wird dann als *meßbar* (bezüglich Φ) bezeichnet.

erfüllt. Es ist klar, daß $C_1 \neq C_0$ ist. Wiederholt man nun für die Menge $C_0 \setminus C_1$ die gleichen Überlegungen wie für C_0, so kommt man zu einer Menge C_2, die die Bedingung

$$\Phi(C_2) \geqq \frac{1}{k_2} \qquad (k_2 > k_1)$$

erfüllt, usw. Setzen wir schließlich

$$F_0 = C_0 \setminus \bigcup_{i=1}^{\infty} C_i,$$

so ist die Menge F_0 nicht leer, weil $\Phi(C_0) < 0$ und $\Phi(C_i) > 0$ für $i \geqq 1$ ist. Aus der Konstruktion von F_0 ergibt sich, daß F_0 negativ ist. Vereinigen wir F_0 mit A^-, so erhalten wir wie oben einen Widerspruch zur Konstruktion von A^-, d. h., für alle meßbaren Mengen $E \subset X \setminus A^-$ gilt

$$\Phi(E) \geqq 0,$$

also ist $X \setminus A^-$ positiv, was zu zeigen war.

Die Zerlegung der Menge X in einen negativen Teil A^- und einen positiven Teil A^+ heißt *Hahnsche Zerlegung.*

Die Hahnsche Zerlegung ist im allgemeinen nicht eindeutig bestimmt. Wenn

$$X = A_1^- \cup A_1^+ \quad \text{und} \quad X = A_2^- \cup A_2^+$$

zwei solche Zerlegungen sind, gilt jedoch für jedes $E \in \mathfrak{S}$

$$\Phi(E \cap A_1^-) = \Phi(E \cap A_2^-) \quad \text{und} \quad \Phi(E \cap A_1^+) = \Phi(E \cap A_2^+). \tag{2}$$

Aus

$$E \cap (A_1^- \setminus A_2^-) \subset E \cap A_1^- \tag{3}$$

folgt nämlich

$$\Phi\big(E \cap (A_1^- \setminus A_2^-)\big) \leqq 0.$$

Andererseits ist auch

$$E \cap (A_1^- \setminus A_2^-) \subset E \cap A_2^+, \tag{4}$$

was

$$\Phi\big(E \cap (A_1^- \setminus A_2^-)\big) \geqq 0$$

nach sich zieht. Also ist

$$\Phi\big(E \cap (A_1^- \setminus A_2^-)\big) = 0.$$

Analog erhält man

$$\Phi\big(E \cap (A_2^- \setminus A_1^-)\big) = 0.$$

Aus den letzten beiden Gleichungen folgt

$$\Phi(E \cap A_1^-) = \Phi(E \cap A_2^-).$$

Genauso zeigt man auch die zweite Gleichung von (2).

Ausgehend von (2) kann man nun feststellen. Das reellwertige Maß Φ bestimmt eindeutig zwei nichtnegative Mengenfunktionen auf $\mathfrak{S}$, nämlich

$$\Phi^+(E) = \Phi(E \cap A^+),$$

$$\Phi^-(E) = -\Phi(E \cap A^-),$$

die die *obere* bzw. die *untere Variation* des reellwertigen Maßes Φ genannt werden und offensichtlich folgende Eigenschaften besitzen:

1. $\Phi = \Phi^+ - \Phi^-$;

2. Φ^+ und Φ^- sind nichtnegative σ-additive Mengenfunktionen, d. h. Maße im üblichen Sinn.

Ein Maß ist offensichtlich auch die Funktion $|\Phi| = \Phi^+ + \Phi^-$; sie wird die *vollständige Variation* des reellwertigen Maßes Φ genannt. Die Darstellung von Φ als Differenz der oberen und der unteren Variation heißt *Jordansche Zerlegung* des reellwertigen Maßes Φ.

Bemerkung. Wir haben bisher nur endliche reellwertige Maße betrachtet, d. h. solche Funktionen Φ, deren Werte nach oben und unten beschränkt waren. Die Funktionen Φ^+ und Φ^- sind in diesem Fall endliche Maße. Alle obigen Betrachtungen können jedoch auch auf reellwertige Maße, die nur nach einer Seite beschränkt sind, ausgedehnt werden.

6.5.2. Grundtypen von reellwertigen Maßen. Es sei μ ein σ-additives Maß, das auf einer σ-Algebra $\mathfrak{S}$ von Untermengen der Menge X definiert ist. Die in $\mathfrak{S}$ enthaltenen Mengen nennen wir *meßbar*. Außerdem werden wir im weiteren folgende Begriffe benutzen.

Wir sagen, daß ein auf $\mathfrak{S}$ definiertes reellwertiges Maß Φ *auf der meßbaren Menge* A_0 *konzentriert* ist, wenn $\Phi(E) = 0$ für jede Menge $E \subset X \setminus A_0$ ist. Die Menge A_0 heißt in diesem Fall *Träger* des reellwertigen Maßes Φ.

Ein reellwertiges Maß Φ heißt *stetig*, wenn $\Phi(E) = 0$ für jede einpunktige Menge E ist. Φ heißt *diskret*, wenn es auf einer endlichen oder abzählbaren Menge konzentriert ist. Diskretheit von Φ bedeutet also die Existenz einer endlichen oder abzählbaren Menge von Punkten $c_1, c_2, \ldots, c_n, \ldots$, so daß für jedes $E \in \mathfrak{S}$

$$\Phi(E) = \sum_{c_k \in E} \Phi(c_k)$$

gilt. Φ heißt *absolut stetig* (bezüglich eines gegebenen Maßes μ), wenn $\Phi(A) = 0$ für jede meßbare Menge A ist, für die $\mu(A) = 0$ ist. Φ heißt *singulär* (bezüglich μ),

wenn Φ auf einer Menge konzentriert ist, die bezüglich μ das Maß 0 hat. Es ist klar, daß ein reellwertiges Maß Φ, das gleichzeitig absolut stetig und singulär bezüglich eines Maßes μ ist, identisch 0 ist.

6.5.3. Absolut stetige reellwertige Maße. Der Satz von Radon-Nikodym. Ein Beispiel für ein reellwertiges Maß, das absolut stetig bezüglich eines gegebenen Maßes μ ist, stellt das als Mengenfunktion betrachtete Lebesguesche Integral

$$\Phi(A) = \int_A f(x)\, d\mu$$

einer festen integrierbaren Funktion f dar. Mit dem folgenden Satz wird bewiesen, daß diese Form bereits alle möglichen Fälle von absolut stetigen reellwertigen Maßen ausschöpft.

Satz 2 (Radon-Nikodym). *Es sei μ ein σ-additives Maß, das auf einer σ-Algebra von Untermengen der Menge X definiert ist, und Φ sei ein reellwertiges Maß, das auf derselben σ-Algebra definiert und absolut stetig bezüglich μ ist. Dann existiert auf X eine bezüglich μ integrierbare Funktion f, so daß*

$$\Phi(A) = \int_A f(x)\, d\mu$$

für jede meßbare Menge A ist. Diese Funktion f, die sogenannte Radon-Nikodym-Ableitung des reellwertigen Maßes Φ nach dem Maß μ, ist bis auf μ-Äquivalenz eindeutig bestimmt.

Beweis. Jedes reellwertige Maß kann nach 6.5.1. als Differenz zweier nichtnegativer Maße dargestellt werden. Ein absolut stetiges Maß wird dabei als Differenz absolut stetiger Maße dargestellt. Daher genügt es, den Beweis des Satzes für nichtnegative reellwertige Maße, d. h. Maße im üblichen Sinn, durchzuführen. Wir beweisen zunächst folgendes Lemma.

Lemma. *Es sei Φ ein bezüglich μ absolut stetiges Maß, das nicht identisch 0 ist. Dann existieren eine Zahl n und eine meßbare Menge B, so daß $\mu(B) > 0$ und B eine positive Menge bezüglich des reellwertigen Maßes $\Phi - \frac{1}{n}\mu$ ist.*

Beweis des Lemmas. $X = A_n^- \cup A_n^+$ sei die Hahnsche Zerlegung für das Maß $\Phi - \frac{1}{n}\mu$, $n = 1, 2, \ldots$, und es sei

$$A_0^- = \bigcap_{n=1}^{\infty} A_n^-, \qquad A_0^+ = \bigcup_{n=1}^{\infty} A_n^+.$$

Dann gilt

$$\Phi(A_0^-) \leqq \frac{1}{n}\mu(A_0^-) \quad \text{für alle } n,$$

d. h. $\Phi(A_0^-) = 0$. Daraus folgt $\Phi(A_0^+) > 0$ und wegen der absoluten Stetigkeit von Φ (bezüglich μ) auch $\mu(A_0^+) > 0$. Damit muß es ein n geben, für das $\mu(A_n^+) > 0$ ist. Diese Zahl n und die Menge $B = A_n^+$ erfüllen die Bedingungen des Lemmas.

Wir gehen nun zum eigentlichen Beweis des Satzes über. Dazu sei K die Menge aller Funktionen f auf X, die die folgenden Eigenschaften besitzen: f ist nichtnegativ, integrierbar bezüglich μ, und es gilt $\int\limits_A f(x)\, d\mu \leqq \Phi(A)$ für jede meßbare Menge A. Weiter sei

$$M = \sup_{f \in K} \left\{ \int\limits_X f(x)\, d\mu \right\},$$

und $\{f_n\}$ sei eine Folge von Funktionen aus K mit der Eigenschaft

$$\lim_{n\to\infty} \int\limits_X f_n(x)\, d\mu = M.$$

Setzen wir nun

$$g_n(x) = \max\big(f_1(x), f_2(x), \ldots, f_n(x)\big),$$

dann ist jede meßbare Menge E in der Form $E = \bigcup\limits_{k=1}^{n} E_k$ darstellbar, wobei $g_n(x) = f_k(x)$ auf E_k und $E_k \cap E_l = 0$ $(k \neq l)$ ist. Daher gilt

$$\int\limits_E g_n(x)\, dx = \sum_{k=1}^{n} \int\limits_{E_k} f_k(x)\, d\mu \leqq \sum_{k=1}^{n} \Phi(E_k) = \Phi(E).$$

Setzen wir

$$f(x) = \sup \{f_n(x)\},$$

so ist offensichtlich $f(x) = \lim\limits_{n\to\infty} g_n(x)$ und folglich nach dem Satz von B. Levi

$$\int\limits_X f(x)\, d\mu = \lim_{n\to\infty} \int\limits_X g_n(x)\, d\mu = M.$$

Wir zeigen nun die Gültigkeit von

$$\Phi(E) - \int\limits_E f(x)\, d\mu = 0.$$

Nach Konstruktion ist die Mengenfunktion

$$\lambda(E) = \Phi(E) - \int\limits_E f(x)\, d\mu$$

nichtnegativ und besitzt alle Eigenschaften eines Maßes. Außerdem ist sie offensichtlich absolut stetig bezüglich μ. Nehmen wir nun an, daß $\lambda \not\equiv 0$ ist, dann gibt es nach dem Lemma eine Zahl $\varepsilon > 0$ und eine Menge B, $\mu(B) > 0$, so daß

$$\varepsilon\mu(E \cap B) \leqq \lambda(E \cap B)$$

für eine beliebige meßbare Menge E ist. Setzen wir $h(x) = f(x) + \varepsilon\chi_B(x)$, wobei $\chi_B(x)$ die charakteristische Funktion der Menge B ist, so erhalten wir für eine beliebige meßbare Menge E

$$\int_E h(x)\,d\mu = \int_E f(x)\,d\mu + \varepsilon\mu(E \cap B) \leqq \int_{E\setminus B} f(x)\,d\mu + \Phi(E \cap B) \leqq \Phi(E),$$

d. h., $h(x)$ gehört einerseits zur oben definierten Menge K, während andererseits

$$\int_X h(x)\,d\mu = \int_X f(x)\,d\mu + \varepsilon\mu(B) > M$$

ist. Aus diesem Widerspruch folgt, daß die Annahme $\lambda \not\equiv 0$ nicht richtig sein kann, also $\lambda \equiv 0$ ist, d. h.

$$\Phi(A) = \int_A f(x)\,d\mu$$

für beliebige meßbare Mengen A gilt. Damit ist die Existenz der im Satz angegebenen Funktion f bewiesen.

Wir zeigen nun die Eindeutigkeit von f. Wenn für alle $A \in \mathfrak{S}$

$$\Phi(A) = \int_A f_1(x)\,d\mu = \int_A f_2(x)\,d\mu$$

angenommen wird, dann folgt für die Mengen

$$A_n = \left\{x\colon [f_2(x) - f_1(x)] > \frac{1}{n}\right\}$$

aus der Čebyševschen Ungleichung (vgl. 5.5.4.)

$$\mu(A_n) \leqq n \int_{A_n} [f_2(x) - f_1(x)]\,d\mu = 0, \qquad n = 1, 2, \ldots$$

Analog ergibt sich für $B_m = \{x\colon [f_1(x) - f_2(x)] > 1/m\}$

$$\mu(B_m) = 0.$$

Wegen

$$\{x\colon f_1(x) \neq f_2(x)\} = \left(\bigcup_n A_n\right) \cup \left(\bigcup_m B_m\right)$$

folgt daraus

$$\mu\{x\colon f_1(x) \neq f_2(x)\} = 0$$

d. h., auf X ist fast überall $f_1(x) = f_2(x)$, was zu zeigen war.

Bemerkung. Der Satz von Radon-Nikodym stellt offensichtlich die natürliche Verallgemeinerung des Satzes von Lebesgue (Satz 3 aus 6.4.) dar, nach dem eine absolut stetige Funktion das Integral ihrer Ableitung ist. Während es jedoch bei der Betrachtung von Funktionen auf der Zahlengeraden eine Methode zur Bestimmung

dieser Ableitung gab (Berechnung des Grenzwertes von $\Delta f/\Delta x$ für $\Delta x \to 0$), stellt der Satz von RADON-NIKODYM nur die Existenz der Ableitung $d\Phi/d\mu$ eines absolut stetigen reellwertigen Maßes Φ nach dem Maß μ fest und liefert keine Möglichkeit zur Berechnung dieser Ableitung. Man kann jedoch eine Methode zur Berechnung dieser Ableitung angeben. Sie besteht darin, daß man den Grenzwert des Quotienten $\Phi(A)/\mu(A)$ für ein System von Mengen berechnet, das sich in einem gewissen Sinn auf den vorgegebenen Punkt zusammenzieht. Wir verzichten hier auf eine Darstellung dieser Methode und verweisen den Leser auf das Buch [61], wo man eine detaillierte Behandlung dieser Fragen findet.

6.6. Das Stieltjessche Integral

6.6.1. Stieltjessche Maße. In 5.1. hatten wir das Lebesguesche Maß auf der Zahlengeraden konstruiert und bereits auf die folgende Möglichkeit der Konstruktion eines allgemeineren Maßes hingewiesen (vgl. 5.1.3.). Man definiert, ausgehend von einer monoton nichtfallenden linksseitig stetigen Funktion F, für alle abgeschlossenen, halboffenen und offenen Teilintervalle des Grundintervalls $[a, b]$ durch die Formeln

$$\left.\begin{aligned} m(\alpha, \beta) &= F(\beta) - F(\alpha + 0), \\ m[\alpha, \beta] &= F(\beta + 0) - F(\alpha), \\ m(\alpha, \beta] &= F(\beta + 0) - F(\alpha + 0), \\ m[\alpha, \beta) &= F(\beta) - F(\alpha) \end{aligned}\right\} \tag{1}$$

ein Maß und setzt dieses Maß mit Hilfe des Lebesgueschen Verfahrens auf eine σ-Algebra $\mathfrak{A}_F$ fort. Diese σ-Algebra enthält alle offenen, alle abgeschlossenen (und damit auch alle Borelschen) Untermengen von $[a, b]$. Die Fortsetzung μ_F des Maßes m auf $\mathfrak{A}_F$, die man auf diese Weise gewinnt, heißt *Lebesgue-Stieltjessches Maß*. Die Funktion F heißt *erzeugende Funktion* dieses Maßes.

Wir betrachten einige Beispiele von Lebesgue-Stieltjesschen Maßen.

1. Ist F eine Sprungfunktion mit den Sprungstellen $x_1, x_2, \ldots$ und den Sprüngen $h_1, h_2, \ldots$ in diesen Punkten, so läßt sich das Maß μ_F, das dieser Funktion entspricht, auf folgende Weise beschreiben: Alle Untermengen des Intervalls $[a, b]$ sind meßbar, und das Maß einer Menge A ist gleich

$$\mu_F(A) = \sum_{x_i \in A} h_i. \tag{2}$$

Aus der Definition des Lebesgue-Stieltjesschen Maßes folgt nämlich sofort, daß das Maß jedes Punktes x_i gleich h_i und das Maß des Komplements der Menge $\{x_i\}_{i=1}^{\infty}$ gleich 0 ist. Daraus folgt auf Grund der σ-Additivität von μ_F die Gleichung (2) für jede Untermenge $A \subset [a, b]$.

Ein Maß μ_F, das einer Sprungfunktion F entspricht, heißt *diskretes* Maß.

2. Ist F eine absolut stetige nichtfallende Funktion auf $[a, b]$, $f = F'$ ihre Ableitung, so ist das Maß μ_F auf allen Lebesgue-meßbaren Untermengen des Intervalls $[a, b]$ definiert, und es gilt für derartige Mengen A

$$\mu_F(A) = \int_A f(x)\,dx. \tag{3}$$

Nach dem Satz von LEBESGUE ist nämlich für jedes offene Intervall (α, β)

$$\mu_F(\alpha, \beta) = F(\beta) - F(\alpha) = \int_\alpha^\beta f(x)\,dx,$$

und da die Lebesguesche Fortsetzung eines σ-additiven Maßes durch ihre Werte auf dem Ausgangssemiring eindeutig bestimmt ist, gilt (3) für beliebige Lebesgue-meßbare Mengen $A \subset [a, b]$.

Ein Maß μ_F, das einer absolut stetigen Funktion F entspricht, heißt *absolut stetiges* Maß.

3. Ist F eine singuläre stetige Funktion, dann entspricht ihr ein Maß μ_F, das ganz auf der Menge (vom Lebesgueschen Maß 0) konzentriert ist, auf der F' verschieden von 0 oder nicht erklärt ist.

Ein Maß μ_F, das einer singulären Funktion F entspricht, heißt *singuläres* Maß.

Es ist klar, daß für eine Funktion $F = F_1 + F_2$ das entsprechende Maß μ_F die Gestalt $\mu_F = \mu_{F_1} + \mu_{F_2}$ hat. Daher folgt aus der Zerlegbarkeit einer monotonen Funktion in eine Summe aus einer Sprungfunktion, einer absolut stetigen und einer singulären Funktion, daß *jedes Lebesgue-Stieltjessche Maß als Summe eines diskreten, eines absolut stetigen und eines singulären Maßes dargestellt werden kann.* Da die Zerlegung der erzeugenden Funktion bis auf konstante Summanden eindeutig ist, ergibt sich entsprechend der Definition (1) für jedes Lebesgue-Stieljessche Maß eine eindeutig bestimmte Zerlegung in eine diskrete, eine absolut stetige und eine singuläre Komponente.

Bisher haben wir Lebesgue-Stieljessche Maße auf einem Intervall betrachtet. Ist jetzt F eine (nach oben und unten) beschränkte monotone nichtfallende Funktion auf der ganzen Geraden und definieren wir analog zu (1) ein Ausgangsmaß für die verschiedenen Intervalle, so erhalten wir durch den Lebesgueschen Fortsetzungsprozeß ein endliches Maß auf der ganzen Geraden, das wir ebenfalls *Lebesgue-Stieltjessches Maß* nennen. Das Maß der ganzen Geraden ist bei diesem Maß gleich

$$F(\infty) - F(-\infty),$$

wobei

$$F(\infty) = \lim_{x\to\infty} F(x), \qquad F(-\infty) = \lim_{x\to-\infty} F(x)$$

ist (die Existenz dieser Grenzwerte folgt aus der Monotonie und Beschränktheit von F).

Der Begriff des Lebesgue-Stieltjesschen Maßes führt nicht, wie es vielleicht auf den ersten Blick erscheint, zu einer speziellen Klasse von Maßen auf der Zahlengeraden, sondern enthält alle möglichen Konstruktionen eines Maßes, d. h., alle endlichen σ-additiven nichtnegativen Mengenfunktionen auf der Geraden sind mit Hilfe dieses Begriffes darstellbar. Denn ist μ ein beliebiges Maß auf der Geraden, so erhält man durch die Festlegung

$$F(x) = \mu(-\infty, x)$$

eine monotone Funktion, deren Lebesgue-Stieljessches Maß mit dem Ausgangsmaß μ übereinstimmt. Auf Grund dieses Sachverhaltes kann man den Begriff des Lebesgue-Stieltjesschen Maßes auf der Geraden als Konstruktionsmethode für Maße auffassen, die eine erzeugende Funktion benutzt.

6.6.2. Das Lebesgue-Stieltjessche Integral. Es sei μ_F ein Maß auf dem Intervall $[a, b]$, das durch die monotone Funktion F erzeugt wird. Für dieses Maß definieren wir wie üblich die Klasse der integrierbaren Funktionen und den Begriff des Lebesgueschen Integrals (vgl. 5.5.)

$$\int_a^b f(x)\, d\mu_F.$$

Ein solches Integral bezüglich eines durch die Funktion F erzeugten Maßes μ_F heißt *Lebesgue-Stieltjessches Integral* und wird mit dem Symbol

$$\int_a^b f(x)\, dF(x)$$

bezeichnet.

Wir betrachten einige Beispiele.

1. Ist F eine Sprungfunktion (d. h. μ_F ein diskretes Maß), dann reduziert sich das Integral

$$\int_a^b f(x)\, dF(x)$$

auf die Summe

$$\sum_i f(x_i)\, h_i,$$

wobei x_i eine Sprungstelle von F und h_i den Sprung von F in x_i bezeichnet.

2. Ist F eine absolut stetige Funktion, dann ist das Lebesgue-Stieljessche Integral

$$\int_a^b f(x)\, dF(x)$$

gleich

$$\int_a^b f(x)\, F'(x)\, dx,$$

d. h. gleich dem Integral über $f(x)\, F'(x)$ bezüglich des gewöhnlichen Lebesgueschen Maßes auf der Geraden.

Das ergibt sich folgendermaßen. Ist $f(x) = \text{const}$, dann folgt

$$\int_a^b f(x)\, dF(x) = \int_a^b f(x)\, F'(x)\, dx \tag{4}$$

aus (3). Auf Grund der σ-Additivität dieser Integrale gilt (4) auch für bezüglich μ_F integrierbare Treppenfunktionen. Ist nun $\{f_n\}$ eine gegen f gleichmäßig konvergente Folge von Treppenfunktionen, die wir o. B. d. A. als monoton nichtfallend voraussetzen können, so konvergiert die ebenfalls monoton nichtfallende Folge $\{f_n(x)\, F'(x)\}$ fast überall gegen $f(x)\, F'(x)$, und nach dem Satz von B. Levi kann in

$$\int_a^b f_n(x)\, dF(x) = \int_a^b f_n(x)\, F'(x)\, dx$$

der Grenzübergang $n \to \infty$ mit der Integralbildung vertauscht werden, woraus die Behauptung folgt.

Aus dem oben über die Zerlegung der Funktion F und des entsprechenden Maßes μ_F Gesagten ergibt sich jetzt, daß bei einer Funktion F, die nur aus einer Sprungfunktion und einer absolut stetigen Funktion besteht, das Lebesgue-Stieltjessche Integral auf eine Reihe (oder endliche Summe) und ein Integral bezüglich des üblichen Lebesgueschen Maßes reduziert werden kann. Wenn jedoch F auch einen singulären Teil enthält, dann ist eine solche Reduktion nicht möglich.

Der Begriff des Lebesgue-Stieltjesschen Integrals kann in natürlicher Weise verallgemeinert werden, indem man von den monotonen Funktionen zu beliebigen Funktionen von beschränkter Variation übergeht. Ist Φ eine solche Funktion, so stellen wir sie als Differenz zweier monotoner Funktionen

$$\Phi = v - g$$

dar (v ist die vollständige Variation der Funktion Φ auf dem Intervall $[a, x]$) und definieren das Lebesgue-Stieltjessche Integral von Φ durch

$$\int_a^b f(x)\, d\Phi(x) = \int_a^b f(x)\, dv(x) - \int_a^b f(x)\, dg(x).$$

Man prüft leicht nach, daß für eine andere Darstellung von Φ als Differenz zweier monotoner Funktionen

$$\Phi = w - h$$

stets

$$\int_a^b f(x)\,dv(x) - \int_a^b f(x)\,dg(x) = \int_a^b f(x)\,dw(x) - \int_a^b f(x)\,dh(x)$$

ist, d. h., zur Berechnung des Lebesgue-Stieltjesschen Integrals einer gegebenen Funktion Φ von beschränkter Variation kann eine beliebige Darstellung dieser Funktion als Differenz zweier monotoner Funktionen benutzt werden.

6.6.3. Einige Anwendungen des Lebesgue-Stieltjesschen Integrals in der Wahrscheinlichkeitsrechnung. Das Lebesgue-Stieltjessche Integral wird sowohl in der Analysis selbst als auch in vielen anderen für die Anwendung wichtigen Disziplinen benutzt. Insbesondere wird dieser Begriff sehr häufig in der Wahrscheinlichkeitsrechnung gebraucht. Wir erinnern daran, daß man eine Funktion F *Verteilungsfunktion* der zufälligen Größe ξ nennt, wenn

$$F(x) = P(\xi < x)$$

gilt, d. h. $F(x)$ die Wahrscheinlichkeit dafür ist, daß die zufällige Größe ξ einen Wert annimmt, der kleiner als x ist. Offensichtlich ist jede Verteilungsfunktion monoton nichtfallend, linksseitig stetig und erfüllt die Bedingungen

$$F(-\infty) = 0, \quad F(+\infty) = 1.$$

Umgekehrt kann man auch jede solche Funktion als Verteilungsfunktion einer zufälligen Größe betrachten.

Wesentliche Charakteristiken einer zufälligen Größe sind ihr Erwartungswert

$$M\xi = \int_{-\infty}^{\infty} x\,dF(x) \tag{5}$$

und ihre Varianz

$$D\xi = \int_{-\infty}^{\infty} (x - M\xi)^2\,dF(x). \tag{6}$$

Die zufälligen Größen teilt man gewöhnlich in diskrete und in stetige zufällige Größen ein. Eine zufällige Größe heißt *diskret*, wenn sie höchstens endlich oder abzählbar viele Werte

$$x_1, x_2, \ldots, x_n, \ldots$$

annehmen kann. So ist z. B. die Anzahl der Telephonanrufe in einer Vermittlungsstation, bezogen auf ein Zeitintervall, eine diskrete zufällige Größe.

Sind $p_1, p_2, \ldots, p_n, \ldots$ die Wahrscheinlichkeiten, mit denen die Größe ξ die Werte $x_1, x_2, \ldots, x_n, \ldots$ annimmt, so ist die Verteilungsfunktion für ξ offensichtlich eine

Sprungfunktion. Für sie reduzieren sich die Integrale (5) und (6) auf die Summen

$$M\xi = \sum_i x_i p_i$$

und

$$D\xi = \sum_i (x_i - a^2)\, p_i \qquad (a = M\xi).$$

Eine zufällige Größe heißt *stetig*, wenn ihre Verteilungsfunktion F absolut stetig ist. Die Ableitung F' dieser Verteilungsfunktion heißt *Dichte der Wahrscheinlichkeitsverteilung* der zufälligen Größe ξ. Entsprechend dem in 6.6.2. Gesagten reduzieren sich für eine stetige zufällige Größe die Lebesgue-Stieltjesschen Integrale ihres Erwartungswertes und ihrer Varianz auf Integrale bezüglich des gewöhnlichen Lebesgueschen Maßes:

$$M\xi = \int_{-\infty}^{\infty} x\, p(x)\, dx,$$

$$D\xi = \int_{-\infty}^{\infty} (x - a)^2\, p(x)\, dx,$$

dabei ist $p = F'$ die Dichte der Wahrscheinlichkeitsverteilung für ξ und $a = M\xi$.

In den elementaren Vorlesungen über Wahrscheinlichkeitsrechnung beschränkt man sich gewöhnlich auf die Betrachtung diskreter und stetiger zufälliger Größen, da man im Grunde genommen bei angewandten Fragen keinen anderen Größen begegnet. Im allgemeinen kann jedoch die Verteilungsfunktion einer zufälligen Größe auch einen singulären Teil enthalten, d. h., nicht jede zufällige Größe kann als Kombination einer diskreten und einer stetigen zufälligen Größe dargestellt werden.

Es sei ξ eine zufällige Größe, F ihre Verteilungsfunktion und $\eta = \varphi(\xi)$ eine andere zufällige Größe, die als Borelfunktion von ξ darstellbar ist. Der Erwartungswert $M\eta$ der Größe η kann entsprechend der Definition durch

$$M\eta = \int_{-\infty}^{\infty} x\, d\Phi(x)$$

beschrieben werden, wobei Φ die Verteilungsfunktion für η ist. Wesentlich ist jedoch, daß dieser Erwartungswert auch mit Hilfe der Verteilungsfunktion F der Größe ξ beschrieben werden kann, wenn die Funktion φ bezüglich des durch F erzeugten Maßes μ_F integrierbar ist. Es gilt dann

$$M\eta = M\varphi(\xi) = \int_{-\infty}^{\infty} \varphi(x)\, dF(x).$$

Um diese Beziehung nachzuweisen, betrachten wir zunächst die durch $y = \varphi(x)$ definierte Abbildung, die die Gerade $-\infty < x < \infty$ mit dem Maß μ_F auf die Gerade $-\infty < y < \infty$ mit dem Maß μ_Φ abbildet. Aus den Resultaten des Kapitels 5 folgt, daß bei einer maßerhaltenden Abbildung φ der Menge X mit dem Maß μ auf die Menge Y mit dem Maß ν für eine auf Y integrierbare Funktion f

$$\int_Y f(y)\, d\nu = \int_X f(\varphi(x))\, d\mu$$

gilt (Substitution der Variablen im Lebesgueschen Integral). Setzen wir in dieser Gleichung $f(y) = y$ und $\mu = \mu_F$, $\nu = \mu_\Phi$, so erhalten wir die zu beweisende Gleichung. Für die Berechnung des Erwartungswertes (und natürlich auch der Varianz) von Funktionen der Größe ξ ist also die Kenntnis der Verteilungsfunktion der Größe ξ ausreichend.

6.6.4. Das Riemann-Stieltjessche Integral. Neben dem Lebesgue-Stieltjesschen Integral, das wir in 6.6.2. als Differenz zweier Lebesguescher Integrale der gegebenen Funktion bezüglich zweier Maße auf der Geraden dargestellt hatten, kann man noch das sogenannte Riemann-Stieltjessche Integral betrachten. Es wird analog zum gewöhnlichen Riemannschen Integral als Grenzwert von Integralsummen definiert.

Es sei Φ wieder eine linksseitig stetige Funktion von beschränkter Variation auf dem halboffenen Intervall $[a, b)$ und f eine beliebige Funktion auf diesem Intervall.[1]) Für eine Zerlegung

$$a = x_0 < x_1 < x_2 < \cdots < x_n = b$$

des Intervalls $[a, b)$ bilden wir die Summe

$$\sum_{i=1}^{n} f(\xi_i)\,[\Phi(x_i) - \Phi(x_{i-1})], \qquad \xi_i \in [x_{i-1}, x_i). \tag{7}$$

Strebt nun diese Summe bei unbegrenzter Verfeinerung der Zerlegung, d. h. für $\max (x_i - x_{i-1}) \to 0$, unabhängig von der Wahl der Teilpunkte x_i und der Punkte ξ_i aus den halboffenen Teilintervallen der entsprechenden Zerlegung gegen einen Grenzwert, dann heißt dieser Grenzwert *Riemann-Stieltjessches Integral* der Funktion f bezüglich der Funktion Φ über $[a, b)$ und wird mit dem Symbol

$$\int_a^b f(x)\, d\Phi(x) \tag{8}$$

bezeichnet.

Satz 1. *Ist die Funktion f stetig auf dem Intervall $[a, b]$, dann existiert ihr Riemann-Stieltjessches Integral* (8) *und stimmt mit dem entsprechenden Lebesgue-Stieltjesschen Integral überein.*

Beweis. Die Summe (7) kann man als Lebesgue-Stieljessches Integral der Treppenfunktion

$$f_n(x) = f(\xi_i) \qquad \text{für } x_{i-1} \leqq x < x_i \qquad (i = 1, \ldots, n)$$

auffassen. Da die Funktion f stetig ist, konvergiert bei unbegrenzter Verfeinerung der Zerlegung des Intervalls $[a, b)$ eine Folge solcher Treppenfunktionen gleichmäßig

---

[1]) Da im Stieltjesschen Integral einzelne Punkte einen von Null verschiedenen Beitrag liefern können, betrachten wir hier halboffene Intervalle.

gegen f. Daraus ergibt sich sofort, daß der Grenzwert der entsprechenden Summen (7) existiert und gleich dem Lebesgue-Stieltjesschen Integral der Grenzfunktion f ist (Grenzwertbildung und Integral können vertauscht werden). Andererseits ist dieser Grenzwert nach Definition gerade das Riemann-Stieltjessche Integral (8) von f. Damit ist der Satz bewiesen.

Wir stellen nun einige elementare Eigenschaften des Riemann-Stieltjesschen Integrals zusammen. Die Funktion f sei dabei immer auf dem ganzen Intervall $[a, b]$ stetig.

1. (*Mittelwertsatz*). *Für das Riemann-Stieltjessche Integral ist folgende Abschätzung richtig:*

$$\left|\int_a^b f(x)\, d\Phi(x)\right| \leqq \max |f(x)| \cdot V_a^b[\Phi]. \tag{9}$$

Dabei bezeichnet $V_a^b[\Phi]$ die vollständige Variation von Φ auf $[a, b]$.

Wir gelangen zu dieser Abschätzung, wenn wir von der für eine beliebige Zerlegung des Intervalls $[a, b)$ gültigen Ungleichung

$$\begin{aligned}\left|\sum_{i=1}^n f(\xi_i)\,[\Phi(x_i) - \Phi(x_{i-1})]\right| &\leqq \sum_{i=1}^n |f(\xi_i)| \cdot |\Phi(x_i) - \Phi(x_{i-1})| \\ &\leqq \max |f(x)| \cdot \sum_{i=1}^n |\Phi(x_i) - \Phi(x_{i-1})| \\ &\leqq \max |f(x)| \cdot V_a^b[\Phi]\end{aligned}$$

ausgehen und den der Integraldefinition entsprechenden Grenzübergang vornehmen.

Im Fall $\Phi(x) = x$ liefert (9) die bekannte Abschätzung

$$\left|\int_a^b f(x)\, dx\right| \leqq (b - a) \max |f(x)|$$

für das Riemannsche Integral.

2. *Ist* $\Phi = \Phi_1 + \Phi_2$, *dann ist*

$$\int_a^b f(x)\, d\Phi(x) = \int_a^b f(x)\, d\Phi_1(x) + \int_a^b f(x)\, d\Phi_2(x).$$

Da für jede Zerlegung des Intervalls $[a, b)$ die entsprechende Gleichung für die Integralsummen besteht, bleibt sie auch beim Grenzübergang erhalten und ist damit auch für die Integrale richtig.

Als wir das Riemann-Stieltjessche Integral (8) definierten, setzten wir voraus, daß die Funktion $\Phi(x)$ linksseitig stetig ist. Die Definition dieses Integrals als Grenzwert der Summen (7) ist jedoch auch für beliebige Funktionen $\Phi(x)$ von beschränkter Variation sinnvoll. Dabei ist dann folgende Eigenschaft von Bedeutung.

3. *Sind Φ_1 und Φ_2 zwei Funktionen von beschränkter Variation auf dem Intervall $[a, b)$, die überall mit Ausnahme von höchstens endlich oder abzählbar vielen inneren Punkten dieses Intervalls übereinstimmen, dann ist für jede auf $[a, b]$ stetige Funktion f*

$$\int_a^b f(x)\, d\Phi_1(x) = \int_a^b f(x)\, d\Phi_2(x).$$

Um diese Aussage zu beweisen, betrachten wir zuerst den Fall $\Phi_2 \equiv 0$, d. h., wir zeigen zunächst die Richtigkeit der folgenden Behauptung.

3'. *Ist ψ eine Funktion von beschränkter Variation, die nur in endlich oder abzählbar vielen Punkten des Intervalls (a, b) von 0 verschieden ist, dann ist für jede auf $[a, b]$ stetige Funktion f*

$$\int_a^b f(x)\, d\psi(x) = 0.$$

Wenn die Funktion ψ nur in einem Punkt x_0 von 0 verschieden ist und Zerlegungen des Intervalls $[a, b)$ betrachtet werden, die x_0 nicht als Teilungspunkt enthalten, dann sind alle Integralsummen gleich 0, und es ergibt sich die obige Gleichung. Auf Grund der Additivität (Eigenschaft 2) ist dann diese Gleichung auch für Funktionen ψ richtig, die in höchstens endlich vielen Punkten von 0 verschieden sind. Ist nun ψ in abzählbar vielen Punkten

$$r_1, r_2, \ldots, r_n, \ldots$$

von 0 verschieden und besitzt dort die Werte

$$y_1, y_2, \ldots, y_n, \ldots,$$

so ist $\sum\limits_n |y_n| < \infty$, da ψ von beschränkter Variation ist. Wir wählen eine Zahl N so, daß $\sum\limits_{n>N} |y_n| < \varepsilon$ ist, und stellen ψ in der Form

$$\psi = \psi_N + \tilde{\psi}$$

dar. Dabei sei ψ_N die Funktion, die in den Punkten $r_1, \ldots, r_N$ die Werte $y_1, \ldots, y_N$ annimmt und sonst gleich 0 ist. Das bedeutet, daß $\tilde{\psi}$ nur in den Punkten $r_{N+1}, r_{N+2}, \ldots$ von 0 verschieden ist. Nach Eigenschaft 2 gilt

$$\int_a^b f(x)\, d\psi(x) = \int_a^b f(x)\, d\psi_N(x) + \int_a^b f(x)\, d\tilde{\psi}(x).$$

Da das erste Integral auf der rechten Seite der Gleichung nach dem bereits Bewiesenen verschwindet, für das zweite Integral aus Eigenschaft 1 zusammen mit der offensichtlichen Ungleichung $V_a{}^b[\tilde{\psi}] = 2 \sum\limits_{n>N} |y_n| < 2\varepsilon$ die Abschätzung

$$\left| \int_a^b f(x)\, d\tilde{\psi}(x) \right| < \max |f(x)| \cdot 2\varepsilon$$

folgt und ε beliebig wählbar ist, ergibt sich die Behauptung 3'.

Um die Eigenschaft 3 zu beweisen, schreiben wir die Funktion Φ_1 in der Form $(\Phi_1 - \Phi_2) + \Phi_2 = \psi + \Phi_2$ und benutzen Eigenschaft 2. Weil ψ nur in endlich oder abzählbar vielen Punkten von 0 verschieden ist, ergibt sich jetzt Eigenschaft 3 aus der bereits bewiesenen Eigenschaft 3'.

Unmittelbar aus Eigenschaft 3 erhält man:

4. *Ist Φ eine Funktion von beschränkter Variation auf $[a, b)$ mit höchstens abzählbar vielen Unstetigkeitsstellen $x_i \in (a, b)$ und ist die Funktion f stetig auf $[a, b]$, dann hängt das Riemann-Stieltjessche Integral $\int_a^b f(x)\, d\Phi(x)$ nicht von den Werten der Funktion Φ in den Unstetigkeitsstellen x_i ab.*

Da das Riemann-Stieltjessche Integral einer stetigen Funktion mit dem entsprechenden Lebesgue-Stieltjesschen Integral übereinstimmt, gilt für das Riemann-Stieltjessche Integral einer stetigen Funktion $f(x)$

$$\int_a^b f(x)\, d\Phi(x) = \sum_i f(x_i)\, h_i,$$

wenn $\Phi(x)$ eine Sprungfunktion ist, und

$$\int_a^b f(x)\, d\Phi(x) = \int_a^b f(x)\, \Phi'(x)\, dx, \tag{10}$$

wenn $\Phi(x)$ eine absolut stetige Funktion ist. Ist dabei die Funktion $\Phi'(x)$ Riemann-integrierbar, so kann das rechte Integral in (10) als gewöhnliches Riemannsches Integral ausgerechnet werden.

Alle Aussagen bezüglich Riemann-Stieltjesscher Integrale über ein endliches Intervall können ohne Schwierigkeiten auf den Fall übertragen werden, daß das Integral über eine Halbgerade oder die ganze Zahlengerade genommen wird.

Bemerkung. Wir haben das Stieltjessche Integral über dem halboffenen Intervall $[a, b)$ betrachtet. Analog kann man dieses Integral über dem Intervall $(a, b]$ oder den Intervallen $[a, b]$ und (a, b) definieren. Im Unterschied zum gewöhnlichen Riemannschen Integral stimmen jedoch die Stieltjesschen Integrale über diesen Intervallen im allgemeinen nicht überein. Ist zum Beispiel a eine Unstetigkeitsstelle der Funktion Φ, dann ist das Integral über $[a, b]$ gleich dem Integral über $(a, b]$ plus dem Term $f(a)\,[\Phi(a + 0) - \Phi(a)]$.

6.6.5. Grenzübergang unter dem Stieltjesschen Integralzeichen. In Kapitel 5 haben wir eine Reihe von Sätzen über den Grenzübergang unter dem Integralzeichen bewiesen. Dabei wurde für eine Folge $\{f_n\}$ von Funktionen gefragt, wann in der Folge ihrer Integrale bei festem Maß der Grenzübergang $n \to \infty$ mit der Integralbildung vertauschbar ist, d. h. wann

$$\lim_{n\to\infty} \left(\int f_n\, d\mu \right) = \int \left(\lim_{n\to\infty} f_n \right) d\mu$$

gilt. Für das Stieltjessche Integral ist nun auch die Frage interessant, wann für eine Folge $\{\Phi_n\}$ maßerzeugender Funktionen von beschränkter Variation in der Folge der Integrale einer festen Funktion f bei variablem Maß Φ_n der Grenzübergang $n \to \infty$ mit der Integralbildung vertauschbar ist, d. h. wann

$$\lim_{n\to\infty} \int_a^b f(x)\, d\Phi_n(x) = \int_a^b f(x)\, d\left(\lim_{n\to\infty} \Phi_n(x)\right)$$

gilt.

Satz 2 (Erster Satz von Helly). *Sind $\Phi_n(x)$ Funktionen von beschränkter Variation auf $[a, b]$, die überall auf $[a, b]$ punktweise gegen eine Funktion $\Phi(x)$ konvergieren und deren vollständige Variationen auf $[a, b]$ beschränkt sind, d. h.*

$$V_a^b[\Phi_n] \leqq C \qquad (n = 1, 2, \ldots),$$

dann ist auch die Grenzfunktion Φ von beschränkter Variation, und es gilt für jede stetige Funktion f

$$\lim_{n\to\infty} \int_a^b f(x)\, d\Phi_n(x) = \int_a^b f(x)\, d\Phi(x). \tag{11}$$

Beweis. Wir zeigen zuerst, daß die vollständige Variation der Grenzfunktion Φ nicht größer ist als die Schranke C der Variationen $V_a^b[\Phi_n]$. Für eine beliebige Zerlegung des Intervalls $[a, b]$ durch die Punkte

$$a = x_0 < x_1 < \cdots < x_m = b$$

ist nämlich

$$\sum_{k=1}^m |\Phi(x_k) - \Phi(x_{k-1})| = \lim_{n\to\infty} \sum_{k=1}^m |\Phi_n(x_k) - \Phi_n(x_{k-1})| \leqq C$$

und deshalb

$$V_a^b[\Phi] \leqq C.$$

Wir beweisen nun die Gleichung (11) zunächst für eine beliebige Treppenfunktion f mit endlichen vielen Werten. Ist $f(x) = h_k$ für $x \in (\alpha_{k-1}, \alpha_k)$, so ist

$$\int_a^b f(x)\, d\Phi_n(x) = \sum_k h_k[\Phi_n(\alpha_k) - \Phi_n(\alpha_{k-1})]$$

und

$$\int_a^b f(x)\, d\Phi(x) = \sum_k h_k[\Phi(\alpha_k) - \Phi(\alpha_{k-1})].$$

Daraus folgt (11), da jeder Summand der ersten Gleichung für $n \to \infty$ in den entsprechenden der zweiten Gleichung übergeht.

Ist schließlich f eine stetige Funktion auf $[a, b]$, dann können wir f für ein beliebiges $\varepsilon > 0$ durch eine Treppenfunktion f_ε mit endlich vielen Werten approximieren, so daß

$$|f(x) - f_\varepsilon(x)| < \frac{\varepsilon}{3C}$$

auf $[a, b]$ gilt. Da

$$\begin{aligned}\left|\int_a^b f(x)\,d\Phi(x) - \int_a^b f(x)\,d\Phi_n(x)\right| \leqq {} & \left|\int_a^b f(x)\,d\Phi(x) - \int_a^b f_\varepsilon(x)\,d\Phi(x)\right| \\ & + \left|\int_a^b f_\varepsilon(x)\,d\Phi(x) - \int_a^b f_\varepsilon(x)\,d\Phi_n(x)\right| \\ & + \left|\int_a^b f_\varepsilon(x)\,d\Phi_n(x) - \int_a^b f(x)\,d\Phi_n(x)\right|\end{aligned}$$

ist und nach dem Mittelwertsatz für das Stieltjessche Integral der erste und der dritte Summand kleiner als $\varepsilon/3$ sind, während der zweite Summand nach den obigen Betrachtungen für hinreichend großes n kleiner als $\varepsilon/3$ wird, folgt hieraus die Behauptung des Satzes.

Bemerkung. Der Satz läßt sich auf den Fall übertragen, daß in den Integralen eine oder beide Grenzen unendlich sind. Dabei muß jedoch von der Funktion f gefordert werden, daß sie im Unendlichen gegen einen endlichen Grenzwert strebt (dann kann f auch auf einem unendlichen Intervall durch Treppenfunktionen mit nur endlich vielen Werten gleichmäßig approximiert werden).

Während der erste Satz von Helly Bedingungen dafür angibt, wann im Riemann-Stieltjesschen Integral der Grenzübergang $n \to \infty$ bei einer Folge $\{\Phi_n\}$ maßerzeugender Funktionen von beschränkter Variation unter das Integralzeichen gezogen werden kann, klärt der zweite Satz von Helly, wann eine Folge von Funktionen $\{\Phi_n\}$ existiert, die den Voraussetzungen des ersten Satzes genügt.

Satz 3 (Zweiter Satz von Helly). *Aus jeder unendlichen Menge M von Funktionen Φ, die auf einem Intervall $[a, b]$ definiert sind und dort den Bedingungen*

$$\max |\Phi(x)| \leqq C, \qquad V_a^b[\Phi] \leqq K \tag{12}$$

(C und K von Φ unabhängige Konstanten) *genügen, kann man eine Folge auswählen, die auf $[a, b]$ überall punktweise konvergiert.*

Beweis. Es genügt, diesen Satz für monotone Funktionen zu beweisen. Denn ist Φ eine beliebige Funktion aus M, so kann Φ als Differenz

$$\Phi = v - g$$

der monotonen Funktionen $v(x) = V_a^x[\Phi]$ und $g(x)$ dargestellt werden, wobei die Funktion v ebenfalls die Bedingungen (12) erfüllt,

$$\max |v(x)| \leqq K, \quad V_a^b[v] = V_a^b[\Phi] \leqq K.$$

Gilt nun die Aussage des Satzes für monotone Funktionen, dann existiert eine Folge $\{\Phi_n\} \subset M$, deren entsprechende Folge $\{v_n\}$ auf $[a, b]$ punktweise gegen eine Grenzfunktion v konvergiert. Da die Funktionen

$$g_n = v_n - \Phi_n$$

ebenfalls monoton sind und den Bedingungen (12) genügen, folgt weiter aus der Aussage des Satzes für monotone Funktionen, daß die Folge $\{\Phi_n\}$ eine Teilfolge $\{\Phi_{n_k}\}$ enthält, deren entsprechende Folge $\{g_{n_k}\}$ auf $[a, b]$ punktweise gegen eine Grenzfunktion g konvergiert. Damit strebt aber dann auf $[a, b]$

$$\Phi_{n_k}(x) = v_{n_k}(x) - g_{n_k}(x) \to v(x) - g(x) = \Phi(x),$$

d. h., aus der Aussage des Satzes für monotone Funktionen folgt die allgemeine Aussage.

Wir beweisen nun den Satz für eine Menge M monotoner Funktionen. Es sei

$$r_1, r_2, \ldots, r_n, \ldots$$

die Folge aller rationalen Zahlen des Intervalls $[a, b]$. Nach (12) ist die Menge aller Zahlen $\Phi(r_1)$ (Φ durchläuft die Menge M) beschränkt; daher gibt es eine Folge $\{\Phi_n^{(1)}\}$ in dieser Menge, die im Punkt r_1 konvergiert. Aus der Folge $\{\Phi_n^{(1)}\}$ wird nun eine Teilfolge $\{\Phi_n^{(2)}\}$ ausgewählt, die auch im Punkt r_2 konvergiert. Aus dieser Folge wählen wir wiederum eine Teilfolge aus, die auch in r_3 konvergiert, usw. Die Diagonalfolge $\{\Phi_n^{(n)}\}$ konvergiert dann in allen rationalen Punkten des Intervalls $[a, b]$. Ihr Grenzwert ist eine nichtfallende Funktion, die bisher nur in den rationalen Punkten $r_1, r_2, \ldots, r_n, \ldots$ erklärt ist. Definieren wir diese Funktion in den übrigen irrationalen Punkten x des Intervalls $[a, b]$ durch $\Phi(x) = \lim\limits_{r \to x-0} \Phi(r)$ (r rational), dann erhalten wir auf diese Weise eine nichtfallende Funktion $\Phi(x)$, die in allen Stetigkeitspunkten Grenzwert der Funktionenfolge $\{\Phi_n^{(n)}\}$ ist. Denn ist x^* ein solcher Punkt, d. h., gibt es für jedes $\varepsilon > 0$ ein $\delta > 0$, so daß

$$|\Phi(x^*) - \Phi(x)| < \frac{\varepsilon}{6} \quad \text{ist, wenn} \quad |x - x^*| < \delta \quad \text{ist,} \tag{13}$$

so kann man zwei rationale Punkte r' und r'', $x^* - \delta < r' < x^* < r'' < x^* + \delta$, auswählen, daß für hinreichend große n, $n > n_0$, die beiden Ungleichungen

$$|\Phi_n^{(n)}(r') - \Phi(r')| < \frac{\varepsilon}{6} \quad \text{und} \quad |\Phi_n^{(n)}(r'') - \Phi(r'')| < \frac{\varepsilon}{6} \tag{14}$$

erfüllt sind. Aus (13) und (14) folgt nun, daß

$$|\Phi_n^{(n)}(r') - \Phi_n^{(n)}(r'')| < \frac{2}{3}\,\varepsilon$$

ist. Da die Funktion $\Phi_n^{(n)}$ nichtfallend ist, gilt außerdem

$$\Phi_n^{(n)}(r') \leqq \Phi_n^{(n)}(x^*) \leqq \Phi_n^{(n)}(r'').$$

Benutzen wir die letzten Ungleichungen bei der Abschätzung des Ausdrucks $|\Phi(x^*) - \Phi_n^{(n)}(x^*)|$, so ergibt sich

$$\begin{aligned} |\Phi(x^*) - \Phi_n^{(n)}(x^*)| &\leqq |\Phi(x^*) - \Phi(r')| + |\Phi(r') - \Phi_n^{(n)}(r')| \\ &\quad + |\Phi_n^{(n)}(r') - \Phi_n^{(n)}(x^*)| \\ &\leqq \frac{\varepsilon}{6} + \frac{\varepsilon}{6} + \frac{4\varepsilon}{6} = \varepsilon \end{aligned}$$

und somit $\lim\limits_{n\to\infty} \Phi_n^{(n)}(x^*) = \Phi(x^*)$.

Bisher wurde also mit der Konstruktion der Folge $\{\Phi_n^{(n)}\}$ eine Folge von Funktionen aus M gefunden, die in allen Stetigkeitspunkten der oben definierten Funktion Φ punktweise gegen Φ konvergiert. Da die Funktion Φ nichtfallend und somit die Menge ihrer Unstetigkeitspunkte höchstens abzählbar ist, kann man durch erneute Anwendung eines Diagonalverfahrens aus der Folge $\{\Phi_n^{(n)}\}$ eine Teilfolge $\{\Phi_n\}$ auswählen, die in allen Punkten des Intervalls $[a, b]$ gegen die Funktionen Φ konvergiert.

Damit ist der Satz vollständig bewiesen.

6.6.6. Die allgemeine Form der linearen stetigen Funktionale im Raum der stetigen Funktionen. In 6.6.3. haben wir schon einige Anwendungen des Stieltjesschen Integrals in der Wahrscheinlichkeitsrechnung betrachtet. Jetzt untersuchen wir noch ein Problem, das diesen Begriff mit der allgemeinen Form eines linearen stetigen Funktionals im Raum $C[a, b]$ in Verbindung bringt.

Satz 4 (F. Riesz). *Jedes lineare stetige Funktional F im Raum $C[a, b]$ ist in der Form*

$$F(f) = \int_a^b f(x)\, d\Phi(x) \tag{15}$$

darstellbar, wobei Φ eine geeignete Funktion von beschränkter Variation ist.[1] *Dabei gilt*

$$\|F\| = V_a^b[\Phi].$$

Beweis. Wird der Raum $C[a, b]$ als Unterraum des Raumes $M[a, b]$ aller auf $[a, b]$ beschränkten Funktionen mit der Norm

$$\|f\| = \sup |f(x)|$$

[1]) Hier wird das Stieltjessche Integral über das Intervall $[a, b]$ erstreckt.

aufgefaßt, so kann nach dem Satz von HAHN-BANACH jedes stetige lineare Funktional über dem Raum $C[a, b]$ unter Erhaltung der Norm auf ganz $M[a, b]$ fortgesetzt werden. Dieses Funktional ist dann insbesondere für alle Funktionen der Form

$$h_\tau(x) = \begin{cases} 1 \text{ für } x \leqq \tau, \\ 0 \text{ für } x > \tau \end{cases}$$

erklärt. Wir setzen

$$\Phi(\tau) = F(h_\tau) \tag{16}$$

und zeigen zuerst, daß diese Funktion auf $[a, b]$ von beschränkter Variation ist. Betrachten wir eine beliebige Zerlegung

$$a = x_0 < x_1 < \cdots < x_n = b$$

dieses Intervalls und setzen

$$\alpha_k = \operatorname{sgn}\big(\Phi(x_k) - \Phi(x_{k-1})\big) \qquad (k = 1, 2, \ldots, n),$$

dann ist

$$\begin{aligned} \sum_{k=1}^{n} |\Phi(x_k) - \Phi(x_{k-1})| &= \sum_{k=1}^{n} \alpha_k\big(\Phi(x_k) - \Phi(x_{k-1})\big) \\ &= \sum_{k=1}^{n} \alpha_k F(h_{x_k} - h_{x_{k-1}}) \\ &= F\left(\sum_{k=1}^{n} \alpha_k (h_{x_k} - h_{x_{k-1}})\right) \\ &\leqq \|F\| \left\|\sum_{k=1}^{n} \alpha_k (h_{x_k} - h_{x_{k-1}})\right\|. \end{aligned}$$

Da die Funktion $\sum_{k=1}^{n} \alpha_k (h_{x_k} - h_{x_{k-1}})$ nur die Werte ± 1 und 0 annimmt, ist ihre Norm gleich 1, und wir erhalten für eine beliebige Zerlegung des Intervalls $[a, b]$

$$\sum_{k=1}^{n} |\Phi(x_k) - \Phi(x_{k-1})| \leqq \|F\|.$$

Daraus folgt

$$V_a^b[\Phi] \leqq \|F\|.$$

Damit haben wir durch den Ansatz (16) aus dem vorgegebenen Funktional F eine Funktion Φ von beschränkter Variation konstruiert. Wir zeigen nun, daß mit Hilfe dieser Funktion das Funktional F in der Form (15) dargestellt werden kann. Es sei f eine beliebige stetige Funktion auf dem Intervall $[a, b]$. Zu einem beliebig vorgegebenen $\varepsilon > 0$ bestimmen wir ein $\delta > 0$, so daß $|f(x'') - f(x')| < \varepsilon$ für $|x'' - x'| < \delta$

ist, zerlegen das Intervall $[a, b]$ durch Teilpunkte x_k in Teilintervalle, die alle eine kleinere Länge als δ haben, und definieren folgende Treppenfunktion:

$$f_\varepsilon(x) = f(x_k) \quad \text{für} \quad x_{k-1} < x \leqq x_k, \quad k = 1, 2, \ldots, n.$$

Diese Funktion kann man offensichtlich auch in der Form

$$f_\varepsilon(x) = \sum_{k=1}^{n} f(x_k) \left[h_{x_k}(x) - h_{x_{k-1}}(x)\right]$$

schreiben, wobei h_τ die in (16) definierte Funktion ist. Nach Konstruktion der Funktion f_ε ist $|f(x) - f_\varepsilon(x)| < \varepsilon$ für alle $x \in [a, b]$, d. h.

$$\|f - f_\varepsilon\| < \varepsilon.$$

Wir bestimmen nun den Wert des Funktionals F für das Element f_ε. Auf Grund der Linearität des Funktionals und der Definition (16) ist

$$F(f_\varepsilon) = \sum_{k=1}^{n} f(x_k) \left[F(h_{x_k}) - F(h_{x_{k-1}})\right] = \sum_{k=1}^{n} f(x_k) \left[\Phi(x_k) - \Phi(x_{k_{k-1}})\right],$$

d. h., $F(f_\varepsilon)$ stimmt mit der der Zerlegung entsprechenden Integralsumme des Integrals

$$\int_a^b f(x)\, d\Phi(x)$$

überein. Daher ist für eine hinreichend feine Zerlegung von $[a, b]$

$$\left| F(f_\varepsilon) - \int_a^b f(x)\, d\Phi(x) \right| < \varepsilon.$$

Andererseits gilt

$$|F(f) - F(f_\varepsilon)| \leqq \|F\| \cdot \|f - f_\varepsilon\| \leqq \|F\| \varepsilon.$$

Daraus ergibt sich

$$\left| F(f) - \int_a^b f(x)\, d\Phi(x) \right| < \varepsilon(1 + \|F\|),$$

was auf Grund der beliebigen Wählbarkeit von ε mit der Gleichung

$$F(f) = \int_b^a f(x)\, d\Phi(x)$$

äquivalent ist.

Wir haben oben gezeigt, daß für die vollständige Variation der in (16) definierten Funktion Φ

$$V_a^b[\Phi] \leqq \|F\| \tag{17}$$

gilt. Nach dem Mittelwertsatz für Riemann-Stieltjessche Integrale folgt aus der obigen Darstellungsformel des Funktionals nun sofort

$$\|F\| \leqq V_a^b[\Phi], \tag{18}$$

und (17) und (18) führen zu

$$\|F\| = V_a^b[\Phi].$$

Damit ist der Satz vollständig bewiesen.

Bemerkung. Es ist klar, daß für eine beliebige Funktion Φ von beschränkter Variation auf $[a, b]$ das Stieltjessche Integral

$$F(f) = \int_a^b f(x)\, d\Phi(x)$$

ein stetiges lineares Funktional auf dem Raum $C[a, b]$ ist. Zwei Funktionen Φ_1 und Φ_2, die auf ganz $[a, b]$ mit Ausnahme höchstens abzählbar vieler innerer Punkte übereinstimmen, erzeugen dabei nach Eigenschaft 3 (vgl. 6.6.4.) ein und dasselbe Funktional. Geht man umgekehrt von zwei Funktionen Φ_1 und Φ_2 aus, für die

$$\int_a^b f(x)\, d\Phi_1(x) = \int_a^b f(x)\, d\Phi_2(x)$$

für jede stetige Funktion f ist, dann ergibt sich sofort $\Phi_1 - \Phi_2 = \text{const}$ in allen Stetigkeitspunkten der Funktion $\Phi_1 - \Phi_2$, d. h. $\Phi_1 - \Phi_2 = \text{const}$ auf dem ganzen Intervall mit Ausnahme von höchstens abzählbar vielen Punkten.

Identifizieren wir in der Klasse der Funktionen von beschränkter Variation auf dem Intervall $[a, b]$ zwei Funktionen, deren Differenz mit Ausnahme von höchstens abzählbar vielen inneren Intervallpunkten konstant ist, so erhalten wir durch Formel (15) eine eineindeutige Zuordnung zwischen stetigen linearen Funktionalen auf $C[a, b]$ und den Funktionen von beschränkter Variation auf $[a, b]$. Jedes lineare Funktional auf $C[a, b]$ kann also als Integral bezüglich eines reellwertigen Maßes geschrieben werden, wobei dieses Maß durch eine Funktion von beschränkter Variation erzeugt wird. Die Zuordnung zwischen den reellwertigen Maßen und den sie erzeugenden Funktionen ist bis auf die oben angegebene Äquivalenz eineindeutig. Dabei gilt für eine beliebige Funktion Φ, die dem Funktional F entspricht, die Ungleichung

$$\|F\| \leqq V_a^b[\Phi].$$

Gleichheit besteht im allgemeinen nicht, doch gibt es, wie der obige Beweis zeigt, in jeder Klasse der bezüglich F äquivalenten Funktionen mindestens eine, für die Gleichheit besteht.

7. Räume summierbarer Funktionen

Eine der wichtigsten Klassen von normierten Räumen ist die Klasse der Räume summierbarer Funktionen. Unter diesen Räumen sind insbesondere der Raum L_1 aller summierbaren Funktionen und der Raum L_2 der quadratisch summierbaren Funktionen von großer Bedeutung. Im weiteren betrachten wir grundlegende Eigenschaften dieser Räume. Dabei stützen wir uns einerseits auf allgemeine Eigenschaften metrischer und linearer normierter Räume, wie sie in den Kapiteln 2 bis 4 dargestellt wurden, andererseits auf den in Kapitel 5 eingeführten Begriff des Lebesgueschen Integrals.

7.1. Der Raum L_1

7.1.1. Definition und grundlegende Eigenschaften des Raumes L_1. Es sei X eine Menge mit dem Maß μ; dabei kann X ein endliches oder auch unendliches Maß haben. Das Maß μ sei vollständig (d. h., jede Untermenge einer beliebigen Menge vom Maß Null sei meßbar).

Wir betrachten die Gesamtheit aller auf X summierbaren Funktionen. Da eine Linearkombination summierbarer Funktionen wieder summierbar ist, bildet diese Gesamtheit bezüglich der üblichen Addition von Funktionen und Multiplikation mit Zahlen einen linearen Raum. Diesen Raum bezeichnen wir mit $L_1(X, \mu)$ oder kürzer mit L_1. Definieren wir für die Funktionen aus L_1 [1])

$$\|f\| = \int |f(x)| \, d\mu, \tag{1}$$

so gilt offensichtlich

$$\|\alpha f\| = |\alpha| \cdot \|f\|$$

und

$$\|f_1 + f_2\| \leqq \|f_1\| + \|f_2\|.$$

Die dritte Eigenschaft einer Norm,

$$\|f\| > 0, \quad \text{wenn} \quad f \neq 0 \quad \text{ist,}$$

---

[1]) Hier und im folgenden wird durch das Symbol $\int$ die Integration über die ganze Menge X gekennzeichnet.

wird erfüllt, wenn wir jeweils alle auf X zueinander äquivalenten Funktionen als nicht verschieden ansehen, d. h. sie als Ausdruck für ein und dasselbe Element aus L_1 auffassen. Insbesondere ist so die Gesamtheit aller Funktionen, die fast überall auf X gleich Null sind, als das Nullelement des Raumes L_1 zu betrachten. Mit dieser Vereinbarung über die zueinander äquivalenten Funktionen besitzt (1) alle Eigenschaften einer Norm, und wir können definieren:

Definition 1. Der normierte Raum aller Klassen zueinander äquivalenter summierbarer Funktionen wird mit L_1 bezeichnet. Dabei ist die Addition von Elementen aus L_1 und ihre Multiplikation mit Zahlen als übliche Addition und Multiplikation von Funktionen[1]) und die Norm durch die Formel

$$\|f\| = \int |f(x)| \, d\mu$$

definiert.

In L_1 wird, wie in jedem normierten Raum, durch

$$\varrho(f, g) = \|f - g\|$$

eine Metrik eingeführt. Die Konvergenz einer Folge summierbarer Funktionen im Sinne dieser Metrik wird als *Konvergenz im Mittel* bezeichnet.

Die Funktionen des Raumes L_1 können komplexwertig (*komplexer Raum* L_1) oder auch nur reellwertig (*reeller Raum* L_1) sein. Die Aussagen dieses Abschnittes sind in beiden Fällen richtig.

Für viele Fragen der Analysis ist die folgende Tatsache von großer Bedeutung.

Satz 1. *Der Raum L_1 ist vollständig.*

Beweis. $\{f_n\}$ sei eine beliebige Fundamentalfolge in L_1, d. h.

$$\|f_n - f_m\| \to 0 \qquad \text{für } n, m \to \infty.$$

Dann kann eine Teilfolge $\{f_{n_k}\}$ mit der Eigenschaft

$$\|f_{n_k} - f_{n_{k+1}}\| = \int |f_{n_k}(x) - f_{n_{k+1}}(x)| \, d\mu < \frac{1}{2^k}$$

ausgewählt werden. Aus dieser Ungleichung und dem Satz von B. Levi folgt, daß die Reihe

$$|f_{n_1}| + |f_{n_2} - f_{n_1}| + \cdots$$

[1]) Genauer: Für zwei Elemente aus L_1, d. h. zwei Klassen zueinander äquivalenter summierbarer Funktionen wird die Summe auf folgende Weise gebildet. Man nimmt aus jeder Klasse einen Repräsentanten (eine Funktion), addiert diese Funktionen im üblichen Sinn und bestimmt die Klasse zueinander äquivalenter summierbarer Funktionen, die die Summe der beiden Repräsentanten enthält. Diese Klasse ist die Summe der vorgegebenen Klassen. Es ist klar, daß das Resultat der Addition von Klassen äquivalenter Funktionen nicht von der Auswahl der Repräsentanten abhängt.
Die Multiplikation von Elementen aus L_1 mit Zahlen wird analog durchgeführt.

fast überall auf X konvergiert. Damit konvergiert aber auch die Reihe

$$f_{n_1} + (f_{n_2} - f_{n_1}) + \cdots$$

fast überall auf X gegen eine Funktion

$$f(x) = \lim_{k\to\infty} f_{n_k}(x),$$

d. h., in L_1 enthält jede Fundamentalfolge eine fast überall konvergente Teilfolge.

Wir zeigen nun, daß die Teilfolge $\{f_{n_k}\}$ sogar im Mittel gegen f konvergiert. Da die Funktionen f_{n_k} Glieder der Fundamentalfolge $\{f_n\}$ sind, gilt für jedes $\varepsilon > 0$

$$\int |f_{n_k}(x) - f_{n_l}(x)|\, d\mu < \varepsilon$$

für alle hinreichend großen k und l. Nach dem Satz von Fatou kann man in dieser Ungleichung den Grenzübergang $l \to \infty$ unter das Integralzeichen ziehen. Wir erhalten

$$\int |f_{n_k}(x) - f(x)|\, d\mu \leqq \varepsilon,$$

d. h., $f \in L_1$ und $f_{n_k} \to f$ in L_1. Wenn aber eine Teilfolge einer Fundamentalfolge konvergiert, dann konvergiert auch die Fundamentalfolge selbst, und zwar gegen denselben Grenzwert. Damit ist der Satz bewiesen.

7.1.2. Überall dichte Mengen in L_1. Nach der Definition des Integrals existiert für jede auf X summierbare Funktion f und jedes $\varepsilon > 0$ eine Treppenfunktion $\varphi(x)$, so daß

$$\int |f(x) - \varphi(x)|\, d\mu < \varepsilon$$

ist. Andererseits ist für eine summierbare Treppenfunktion mit den Werten $y_1, y_2, \ldots$ auf den Mengen $E_1, E_2, \ldots$ das Integral gleich der Summe der absolut konvergenten Reihe

$$\sum_{n=1}^{\infty} y_n \mu(E_n).$$

Daraus folgt, daß jede summierbare Treppenfunktion als Grenzwert (im Mittel) einer Folge von Treppenfunktionen mit nur endlich vielen Werten dargestellt werden kann. Zusammen mit der ersten Aussage heißt das: *Die Funktionen mit nur endlich vielen Werten (d. h. die endlichen Linearkombinationen charakteristischer Funktionen) sind im Raum L_1 überall dicht.*

Es sei R ein metrischer Raum und μ ein Maß auf R mit der Eigenschaft, daß alle offenen und abgeschlossenen Mengen in R meßbar sind und für jede meßbare Menge $M \subset R$

$$\mu(M) = \inf_{M \subset G} \mu(G) \tag{2}$$

ist, wobei das Infimum über alle offenen Mengen $G \supset M$ gebildet wird. Eine solche Bedingung ist im Fall des Lebesgueschen Maßes auf dem euklidischen Raum und in

vielen anderen praktisch interessanten Fällen erfüllt. Unter diesen Voraussetzungen gilt folgender Satz.

Satz 2. *Die Menge aller stetigen Funktionen ist überall dicht in* $L_1(R, \mu)$.

Beweis. Nach den vorangehenden Überlegungen genügt es zu zeigen, daß jede Treppenfunktion mit nur endlich vielen Werten als Grenzwert (im Mittel) einer Folge stetiger Funktionen dargestellt werden kann. Da jede Treppenfunktion mit nur endlich vielen Werten eine Linearkombination von charakteristischen Funktionen $\chi_M(x)$ meßbarer Mengen ist, genügt es, die letzteren durch stetige Funktionen im Mittel zu approximieren. Ist M eine meßbare Menge im metrischen Raum R, dann folgt aus der Voraussetzung (2), daß für jedes $\varepsilon > 0$ eine abgeschlossene Menge F_M und eine offene Menge G_M mit den Eigenschaften

$$F_M \subset M \subset G_M \quad \text{und} \quad \mu(G_M) - \mu(F_M) < \varepsilon$$

gefunden werden können. Setzen wir nun [1])

$$\varphi_\varepsilon(x) = \frac{\varrho(x, R \setminus G_M)}{\varrho(x, R \setminus G_M) + \varrho(x, F_M)},$$

so ist die Funktion gleich 0 für $x \in R \setminus G_M$, gleich 1 für $x \in F_M$ und stetig auf R, da jede der Funktionen $\varrho(x, R \setminus G_M)$ und $\varrho(x, F_M)$ stetig ist und die Summe dieser Funktionen nirgends verschwindet. Die Funktion $|\chi_M - \varphi_\varepsilon|$ ist dann auf $G_M \setminus F_M$ nicht größer als 1 und außerhalb dieser Menge gleich 0. Damit ergibt sich

$$\int |\chi_M(x) - \varphi_\varepsilon(x)| \, d\mu < \varepsilon,$$

was zu zeigen war.

Der Raum $L_1(X, \mu)$ hängt offenbar von der Auswahl der Menge X und des Maßes μ auf ihr ab. Ist z. B. das Maß μ in endlich vielen Punkten konzentriert, so ist $L_1(X, \mu)$ ein endlichdimensionaler Raum. Eine grundlegende Rolle spielen in der Analysis unendlichdimensionale Räume $L_1(X, \mu)$, in denen es eine abzählbare überall dichte Untermenge gibt. Um solche Räume charakterisieren zu können, führen wir noch einen Begriff ein, der eigentlich zur allgemeinen Maßtheorie gehört.

Definition 2. Ein Maß μ heißt *Maß mit abzählbarer Basis*, wenn es ein abzählbares System

$$\mathscr{A} = \{A_n\} \qquad (n = 1, 2, \ldots)$$

meßbarer Mengen $A_n \subset X$ (*abzählbare Basis* des Maßes μ) gibt, so daß für jede meßbare Menge $M \subset X$ und jedes $\varepsilon > 0$ eine Menge $A_k \in \mathscr{A}$ mit der Eigenschaft

$$\mu(M \,\triangle\, A_k) < \varepsilon$$

gefunden werden kann.

[1]) $\varrho(x, A)$ bezeichnet den Abstand des Punktes x von der Menge A.

Insbesondere besitzt das Maß μ eine abzählbare Basis, wenn man es als Lebesguesche Fortsetzung eines Maßes m, das auf einem abzählbaren Semiring $\mathfrak{S}$ definiert ist, auffassen kann. Denn in diesem Fall ist der Ring $\mathfrak{R}(\mathfrak{S})$ offensichtlich ebenfalls abzählbar und eine Basis für das Maß μ. Hieraus folgt z. B., daß das Lebesguesche Maß auf einem Intervall eine abzählbare Basis besitzt, da man als Ausgangssemiring die Gesamtheit aller halboffenen Intervalle mit rationalen Endpunkten nehmen kann.

Das Produkt $\mu = \mu_1 \times \mu_2$ zweier Maße mit abzählbarer Basis besitzt wieder eine abzählbare Basis. Eine solche Basis bilden z. B. alle endlichen Vereinigungen der Produkte $A_k^{(1)} \times A_l^{(2)}$, deren erster Faktor aus der Basis von μ_1 und deren zweiter Faktor aus der Basis von μ_2 stammt. Daher besitzt das Lebesguesche Maß in der Ebene (und entsprechend auch im n-dimensionalen Raum) eine abzählbare Basis.

Satz 3. *Wenn das Maß μ eine abzählbare Basis besitzt, gibt es in $L_1(X, \mu)$ eine abzählbare überall dichte Menge von Funktionen*

$$f_1, f_2, \ldots, f_n, \ldots.$$

Beweis. Wir zeigen, daß die endlichen Summen

$$\sum_{k=1}^{n} c_k f_k(x), \tag{3}$$

in denen die c_k rationale Zahlen und die Funktionen f_k die charakteristischen Funktionen der Elemente der Basis des Maßes μ sind, eine abzählbare überall dichte Menge in $L_1(X, \mu)$ bilden.

Früher wurde gezeigt, daß die Menge der Treppenfunktionen mit endlich vielen Werten in L_1 überall dicht ist. Da jede solche Treppenfunktion beliebig gut durch Treppenfunktionen mit nur rationalen Werten approximiert werden kann, genügt es für den Beweis des Satzes zu zeigen, daß eine beliebige Treppenfunktion f, die die Werte

$$y_1, y_2, \ldots, y_n \qquad (\text{alle } y_i \text{ rational})$$

auf den Mengen

$$E_1, E_2, \ldots, E_n \qquad \left(\bigcup_{i=1}^{n} E_i = X,\ E_i \cap E_j = \emptyset \quad \text{für} \quad i \neq j\right)$$

annimmt, beliebig gut im Sinne der Metrik des Raumes L_1 durch Funktionen der Gestalt (3) approximiert werden kann. Die Abzählbarkeit dieser Menge ist evident.

Das Mengensystem

$$A_1^*, A_2^*, \ldots, A_n^*, \ldots \tag{4}$$

sei eine abzählbare Basis für das Maß μ. Man sieht sofort, daß dieses System zu einem Ring

$$A_1, A_2, \ldots, A_n, \ldots \tag{5}$$

erweitert werden kann, der ebenfalls eine abzählbare Basis für das Maß μ ist. Damit gibt es für jedes $\varepsilon > 0$ Mengen A_{n_k} aus dem System (5), so daß

$$\mu(E_k \,\triangle\, A_{n_k}) < \varepsilon$$

ist. Setzen wir

$$A_k' = A_{n_k} \setminus \bigcup_{i<k} A_{n_i} \qquad (k = 1, 2, \ldots, n)$$

und definieren

$$f^*(x) = \begin{cases} y_k & \text{für} \quad x \in A_k', \\ 0 & \text{für} \quad x \in X \setminus \bigcup_{i=1}^{n} A_i', \end{cases}$$

so sieht man leicht, daß für genügend kleines $\varepsilon > 0$ das Maß

$$\mu\{x: f(x) \neq f^*(x)\}$$

beliebig klein gemacht werden kann. Damit wird auch das Integral

$$\int |f(x) - f^*(x)| \, d\mu \leqq \left(2 \max_k |y_k|\right) \mu\{x: f(x) \neq f^*(x)\}$$

für hinreichend kleines $\varepsilon > 0$ beliebig klein. Auf Grund der Ringeigenschaft der Basis (5) ist $f^*(x)$ eine Funktion von der Form (3), was zu zeigen war.

Ist X ein Intervall der Zahlengeraden und μ das Lebesguesche Maß, dann ist z. B. die Menge aller Polynome mit rationalen Koeffizienten eine abzählbare überall dichte Untermenge in $L_1(X, \mu)$. Denn diese Menge ist offensichtlich abzählbar und nach dem Satz von Weierstrass bezüglich der gleichmäßigen Konvergenz überall dicht in der Menge der stetigen Funktionen, während letztere nach Satz 2 eine überall dichte Menge in $L_1(X, \mu)$ bilden.

7.2. Der Raum L_2

7.2.1. Definition und grundlegende Eigenschaften des Raumes L_2. Wie wir im letzten Abschnitt sahen, ist der Raum L_1 ein vollständiger linearer normierter Raum (Banachraum). Im Sinne von Satz 8 aus 3.4. („Parallelogrammsatz") kann seine Norm jedoch nicht durch ein Skalarprodukt erzeugt werden. Denn für die auf $[0, 2\pi]$ integrierbaren Funktionen $f \equiv 1$, $g = \sin x$ ist in der Norm von L_1 die Beziehung

$$\|f + g\|^2 + \|f - g\|^2 = 2(\|f\|^2 + \|g\|^2)$$

nicht erfüllt, d. h., der Raum L_1 ist nicht unitär.

Einen Funktionenraum, der nicht nur normiert, sondern auch unitär ist, erhält man, wenn man von der Gesamtheit aller Funktionen mit integrierbarem Quadrat ausgeht. Zunächst betrachten wir nur reellwertige Funktionen f, die auf einer

Menge X mit dem Maß μ definiert sind.[1]) Alle Funktionen werden als meßbar vorausgesetzt, äquivalente Funktionen werden als nicht verschieden betrachtet.

Definition 1. Eine Funktion f heißt *Funktion mit integrierbarem Quadrat* (oder *quadratisch integrierbar*) *auf* X, wenn das Integral

$$\int f^2(x)\, d\mu$$

existiert und einen endlichen Wert besitzt. Die Gesamtheit aller solcher Funktionen bezeichnen wir mit $L_2(X, \mu)$ oder kürzer mit L_2.

Die quadratisch integrierbaren Funktionen besitzen die folgenden grundlegenden Eigenschaften.

1. *Das Produkt zweier quadratisch integrierbarer Funktionen ist eine integrierbare Funktion.*

Das folgt unmittelbar aus der Ungleichung

$$|f(x)\, g(x)| \leqq \frac{1}{2}\, [f^2(x) + g^2(x)]$$

und den Eigenschaften des Lebesgueschen Integrals.

Folgerung. *Jede auf einer Menge mit endlichem Maß quadratisch integrierbare Funktion f ist dort auch integrierbar.*

Das ergibt sich aus Eigenschaft 1, wenn dort $g(x) \equiv 1$ gesetzt wird.

2. *Die Summe zweier Funktionen aus L_2 gehört ebenfalls zu L_2.*

Es gilt nämlich

$$\big(f(x) + g(x)\big)^2 \leqq f^2(x) + 2\, |f(x)\, g(x)| + g^2(x),$$

und nach der Voraussetzung sowie Eigenschaft 1 sind alle drei Funktionen der rechten Seite integrierbar.

3. *Wenn $f \in L_2$ und α eine beliebige reelle Zahl ist, dann ist $\alpha f \in L_2$.*

Für $f \in L_2$ gilt nämlich

$$\int [\alpha f(x)]^2\, d\mu = \alpha^2 \int f^2(x)\, d\mu < \infty.$$

Aus den Eigenschaften 2 und 3 folgt, daß eine Linearkombination von Funktionen aus L_2 ebenfalls zu L_2 gehört. Dabei erfüllt die Addition der Funktionen aus L_2 und ihre Multiplikation mit reellen Zahlen offensichtlich alle Bedingungen, die in der

---

[1]) Anm. d. Übers.: In 7.2.4. werden mit derselben Zielstellung komplexwertige Funktionen betrachtet. Da eine weitgehende Analogie besteht, werden dort nur die Grunddefinitionen und die Hauptresultate formuliert. Im Hinblick auf spätere Untersuchungen, bei denen hauptsächlich im komplexen Raum L_2 gerechnet wird, ist es ratsam, ausgehend von den entsprechenden Definitionen in 7.2.4., zur Übung auch die Rechnungen aus 7.2.1. und 7.2.2. auf den komplexen Fall zu übertragen.

Definition des linearen Raumes (vgl. 3.1.) genannt wurden, so daß festgestellt werden kann: *Die Gesamtheit der quadratisch integrierbaren Funktionen bildet einen linearen Raum.*

Setzen wir nun für Funktionen $f, g \in L_2$

$$(f, g) = \int f(x)\, g(x)\, d\mu,$$

dann erfüllt dieser Ansatz alle Forderungen, die in der Definition des Skalarprodukts (vgl. 3.4.) genannt wurden. So ist offensichtlich:

1. $(f, g) = (g, f)$,
2. $(f_1 + f_2, g) = (f_1, g) + (f_2, g)$,
3. $(\alpha f, g) = \alpha(f, g)$,
4. $(f, f) > 0$, wenn $f \neq 0$,

weil wir vereinbart hatten, zueinander äquivalente Funktionen nicht zu unterscheiden (d. h. die Gesamtheit aller auf X zur Funktion $f \equiv 0$ äquivalenten Funktionen als das Nullelement in L_2 zu betrachten).

Nachdem wir für die quadratisch integrierbaren Funktionen eine Addition, eine Multiplikation mit reellen Zahlen und ein Skalarprodukt eingeführt haben, können wir abschließend definieren:

Definition 2. Der lineare Raum aller Klassen zueinander äquivalenter quadratisch integrierbarer Funktionen, in welchem durch

$$(f, g) = \int f(x)\, g(x)\, d\mu$$

ein Skalarprodukt eingeführt ist, heißt *unitärer Raum* L_2.

In L_2 ist, wie in jedem unitären Raum, die Schwarzsche Ungleichung und die Dreiecksungleichung erfüllt. Diese Ungleichungen haben hier die Form

$$\left(\int f(x)\, g(x)\, d\mu\right)^2 \leqq \int f^2(x)\, d\mu \cdot \int g^2(x)\, d\mu$$

und

$$\sqrt{\int (f(x) + g(x))^2\, d\mu} \leqq \sqrt{\int f^2(x)\, d\mu} + \sqrt{\int g^2(x)\, d\mu}.$$

Für den Fall $\mu(X) < \infty$ und $g(x) \equiv 1$ liefert die Schwarzsche Ungleichung folgende nützliche Abschätzung:

$$\left(\int f(x)\, d\mu\right)^2 \leqq \mu(X) \int f^2(x)\, d\mu. \tag{1}$$

Die Norm in L_2 wird durch

$$\|f\| = \sqrt{(f, f)} = \sqrt{\int f^2(x)\, d\mu},$$

die Metrik in L_2 durch

$$\varrho(f, g) = \|f - g\| = \sqrt{\int (f(x) - g(x))^2\, d\mu}$$

definiert. Die Größe

$$\int \bigl(f(x) - g(x)\bigr)^2 \, d\mu = \|f - g\|^2$$

wird auch als *mittlere quadratische Abweichung* der Funktionen f und g voneinander bezeichnet.

Die Konvergenz einer Funktionenfolge im Sinne der Metrik des Raumes L_2 wird als *Konvergenz im Quadratmittel* bezeichnet. Wenn keine Verwechslung mit der in 7.1. betrachteten Konvergenz in L_1 zu befürchten ist, werden wir auch für die Konvergenz in L_2 die kürzere Bezeichnung „Konvergenz im Mittel“ verwenden.

Satz 1. *Der Raum $L_2(X, \mu)$ ist für $\mu(X) < \infty$ vollständig.*

Beweis. Es sei $\{f_n\}$ eine beliebige Fundamentalfolge in L_2, d. h.

$$\|f_n - f_m\| \to 0 \qquad \text{für} \quad n, m \to \infty .$$

Dann ist auf Grund der Abschätzung (1)

$$\begin{aligned} \int |f_n(x) - f_m(x)| \, d\mu &\leqq [\mu(X)]^{1/2} \Bigl\{ \int \bigl(f_n(x) - f_m(x)\bigr)^2 \, d\mu \Bigr\}^{1/2} \\ &\leqq \varepsilon \, [\mu(X)]^{1/2}, \end{aligned} \tag{2}$$

d. h., $\{f_n\}$ ist auch Fundamentalfolge in der Metrik des Raumes L_1. Wiederholen wir die Überlegungen, die beim Beweis der Vollständigkeit des Raumes L_1 angestellt wurden, so ergibt sich auch hier: $\{f_n\}$ enthält eine Teilfolge $\{f_{n_k}\}$, die fast überall gegen eine Funktion f konvergiert. Aus der für alle genügend großen k und l gültigen Ungleichung

$$\int \bigl(f_{n_k}(x) - f_{n_l}(x)\bigr)^2 \, d\mu < \varepsilon$$

ergibt sich dann durch Grenzübergang $l \to \infty$ unter Verwendung des Satzes von Fatou

$$\int \bigl(f_{n_k}(x) - f(x)\bigr)^2 \, d\mu \leqq \varepsilon,$$

d. h., $f \in L_2$ und $f_{n_k} \to f$ in L_2. Aus der Tatsache, daß eine Teilfolge jeder Fundamentalfolge im Raum L_2 konvergiert, folgt nun wieder, wie beim Beweis von Satz 1 aus 7.1., daß die Fundamentalfolge selbst in L_2 konvergiert. Damit ist der Satz bewiesen.

7.2.2. Eigenschaften quadratisch integrierbarer Funktionen über Mengen unendlichen Maßes. In den vorhergehenden Betrachtungen wurden manche Eigenschaften nur für quadratisch integrierbare Funktionen bewiesen, die auf einer Menge X endlichen Maßes definiert waren. Die Bedingung $\mu(X) < \infty$ wurde dabei wesentlich benutzt. So war diese Bedingung Voraussetzung für die Aussage, daß jede quadratisch integrierbare Funktion auch integrierbar ist, und Voraussetzung für die Ungleichung (2), auf die sich der Beweis der Vollständigkeit des Raumes L_2 stützte. Wenn wir Funktionen auf einer Menge mit unendlichem Maß betrachten (z. B. auf der ganzen

Zahlengeraden mit dem Lebesgueschen Maß), dann gehört nicht jede Funktion aus L_2 auch zu L_1. Zum Beispiel ist die Funktion $\frac{1}{\sqrt{1+x^2}}$ auf der ganzen Zahlengeraden nicht integrierbar, während ihr Quadrat integrierbar ist.

Weiter folgte im Fall $\mu(X) < \infty$ mit Hilfe von (1) aus der Konvergenz einer Funktionenfolge in L_2 die Konvergenz dieser Folge in L_1. Für $\mu(X) = \infty$ ist diese Folgerung nicht richtig. So konvergiert z. B. die Folge

$$\varphi_n(x) = \begin{cases} \frac{1}{n} & \text{für} \quad |x| \leqq n, \\ 0 & \text{für} \quad |x| > n \end{cases}$$

gegen 0 im Raum $L_2(-\infty, \infty)$ der quadratisch integrierbaren Funktionen auf der Zahlengeraden, während sie in $L_1(-\infty, \infty)$ keinen Grenzwert besitzt.

Für die bisher aufgezählten Folgerungen ist also die Bedingung $\mu(X) < \infty$ wesentlich, für die Vollständigkeit des Raumes L_2 ist sie es nicht, d. h., *Satz* 1 *bleibt im Fall* $\mu(X) = \infty$[1]) *richtig.*

Wir beweisen nun diese Behauptung. Wie bei der Einführung des Integrals über eine Menge unendlichen Maßes (vgl. 5.5.6.) setzen wir voraus, daß die ganze Menge X als abzählbare Vereinigung von Mengen endlichen Maßes dargestellt werden kann.

Es sei

$$X = \bigcup_{n=1}^{\infty} X_n, \qquad \mu(X_n) < \infty, X_n \cap X_m = \emptyset \text{ für } n \neq m,$$

eine solche Darstellung und $\{f_n\}$ eine Fundamentalfolge in $L_2(X, \mu)$. Dann gibt es für jedes $\varepsilon > 0$ eine Zahl N, so daß

$$\int [f_k(x) - f_l(x)]^2 \, d\mu < \varepsilon \qquad \text{für alle } k, l \geqq N$$

ist. Setzen wir

$$f_k^{(n)}(x) = \begin{cases} f_k(x) & \text{für } x \in X_n, \\ 0 & \text{sonst}, \end{cases}$$

so gilt auf Grund der σ-Additivität des Lebesgueschen Integrals

$$\int [f_k(x) - f_l(x)]^2 \, d\mu = \sum_{n=1}^{\infty} \int_{X_n} [f_k^{(n)}(x) - f_l^{(n)}(x)]^2 \, d\mu < \varepsilon.$$

Damit ist für jedes endliche M erst recht

$$\sum_{n=1}^{M} \int_{X_n} [f_k^{(n)}(x) - f_l^{(n)}(x)]^2 \, d\mu < \varepsilon. \tag{3}$$

---

[1]) Der in 7.1. angegebene Beweis für die Vollständigkeit des Raumes L_1 benötigt keine Voraussetzungen über die Endlichkeit des Maßes der Menge X.

Da die Gesamtheit der quadratisch integrierbaren Funktionen auf jedem X_n ein vollständiger Raum ist, gibt es für jedes n eine Grenzfunktion

$$f^{(n)}(x) = \lim_{l \to \infty} f_l^{(n)}(x)$$

im Sinne der Konvergenz des Raumes $L_2(X_n, \mu)$. Führen wir nun in (3) den Grenzübergang $l \to \infty$ durch und vertauschen Grenzübergang und Integral entsprechend dem Satz von FATOU, so erhalten wir

$$\sum_{n=1}^{M} \int_{X_n} [f_k^{(n)}(x) - f^{(n)}(x)]^2 \, d\mu \leqq \varepsilon$$

für jedes M. Der Grenzübergang $M \to \infty$ liefert nun

$$\sum_{n=1}^{\infty} \int_{X_n} [f_k^{(n)}(x) - f^{(n)}(x)]^2 \, d\mu \leqq \varepsilon .$$

Setzen wir

$$f(x) = f^{(n)}(x) \qquad \text{für } x \in X_n,$$

dann kann die letzte Ungleichung in der Form

$$\int [f_k(x) - f(x)]^2 \, d\mu \leqq \varepsilon$$

geschrieben werden. Hieraus folgt, daß $f(x)$ zu $L_2(X, \mu)$ gehört und daß die Folge $f_k(x)$ im Sinne der Konvergenz des Raumes $L_2(X, \mu)$ gegen $f(x)$ konvergiert. Damit ist die Vollständigkeit des Raumes $L_2(X, \mu)$ auch im Fall $\mu(X) = \infty$ bewiesen.

Aufgabe. $L_p(X, \mu)$ sei die Gesamtheit aller Klassen zueinander äquivalenter Funktionen f, für die $\int |f(x)|^p \, d\mu < \infty$ ist ($1 \leqq p < \infty$). Man zeige, daß $L_p(X, \mu)$ ein Banachraum bezüglich der Norm $\|f\| = (\int |f(x)|^p \, d\mu)^{1/p}$ ist.

7.2.3. Überall dichte Mengen in L_2. Isomorphie. Bisher haben wir festgestellt, daß der Raum $L_2(X, \mu)$ der quadratisch integrierbaren Funktionen ein vollständiger unitärer Raum ist. Mit Ausnahme der entarteten Fälle ist dieser Raum unendlichdimensional. Daher ist es für verschiedene Anwendungen wichtig zu klären, wann der Raum $L_2(X, \mu)$ separabel ist, d. h. eine abzählbare überall dichte Menge enthält. In 7.1. hatten wir gezeigt, daß für den Raum $L_1(X, \mu)$ die Separabilität aus der Existenz einer abzählbaren Basis für das Maß μ folgt. Man überzeugt sich nun leicht davon, daß diese Bedingung auch die Separabilität von $L_2(X, \mu)$ garantiert. Da jede Funktion aus $L_2(X, \mu)$ beliebig gut durch Funktionen approximiert werden kann, die außerhalb einer Menge endlichen Maßes verschwinden[1]), genügt es, in der Gesamtheit dieser Funktionen eine abzählbare überall dichte Menge auszuwählen. Die gleichen Überlegungen wie beim Beweis des Satzes 3 aus 7.1. zeigen, daß dies unter

[1]) Wenn $\mu(X) < \infty$ ist, entfällt diese Überlegung.

der angegebenen Voraussetzung immer möglich ist. Besitzt also das Maß μ eine abzählbare Basis, dann ist der Raum $L_2(X, \mu)$ ein vollständiger separabler unitärer Raum. Schließt man den Fall endlichdimensionaler Räume aus, so kann man sagen: *Wenn das Maß μ eine abzählbare Basis besitzt, dann ist $L_2(X, \mu)$ ein separabler Hilbertraum.*

Nach dem Satz über die Isomorphie separabler Hilberträume sind alle solche Räume $L_2(X, \mu)$ zueinander isomorph. Insbesondere ist jeder solche Raum $L_2(X, \mu)$ isomorph zum Raum l_2 der quadratisch summierbaren Zahlenfolgen. Dieser Raum l_2 kann selbst auch als spezieller Raum $L_2(X, \mu)$ aufgefaßt werden, indem man für X eine abzählbare Menge und für μ das σ-additive Maß wählt, das auf allen Untermengen von X definiert und für alle einpunktigen Untermengen von X gleich 1 ist.

Im weiteren werden wir nur Räume $L_2(X, \mu)$ betrachten, bei denen das Maß μ eine abzählbare Basis besitzt. Wenn keine Mißverständnisse zu befürchten sind, werden wir solche Räume einfach mit L_2 bezeichnen.

Da der Raum L_2 eine Realisierung eines abstrakten Hilbertraumes darstellt, können alle im abstrakten Fall eingeführten Begriffe und bewiesenen Aussagen (vgl. 3.4.) übertragen werden.

Insbesondere folgt so aus dem Satz von Riesz, nach dem jedes lineare stetige Funktional F auf dem Hilbertraum H in Form eines Skalarprodukts

$$F(h) = (h, a)$$

mit einem festen Element $a \in H$ geschrieben werden kann: *Jedes lineare stetige Funktional auf L_2 hat die Form*

$$F(f) = \int f(x)\, g(x)\, d\mu,$$

wobei g eine feste auf X quadratisch integrierbare Funktion ist.

7.2.4. Der komplexe Raum L_2. Bisher haben wir nur den reellen Raum L_2 betrachtet. Alle dabei gewonnenen Aussagen lassen sich leicht auf den komplexen Fall übertragen. Wir nennen eine komplexwertige Funktion f, die auf einer Menge X mit dem Maß μ definiert ist, *Funktion mit integrierbarem Quadrat* (oder *quadratisch integrierbar*), wenn das Integral

$$\int_X |f(x)|^2\, d\mu$$

existiert und endlich ist. Definieren wir die Summe zweier solcher Funktionen und ihre Multiplikation mit komplexen Zahlen in der üblichen Weise sowie das Skalarprodukt durch die Formel

$$(f, g) = \int_X f(x)\, \overline{g(x)}\, d\mu,$$

so erhalten wir einen unitären Raum, der *komplexer Raum L_2* genannt wird (dabei sehen wir, wie im reellen Fall, zueinander äquivalente Funktionen als ein und dasselbe

Element des Raumes an). Dieser Raum ist vollständig und, wenn das Maß μ eine abzählbare Basis besitzt, auch separabel. Schließt man endlichdimensionale Räume aus, so kann man sagen: *Wenn das Maß μ eine abzählbare Basis besitzt, ist der komplexe Raum $L_2(X, \mu)$ ein separabler Hilbertraum.* Alle solche Räume sind zueinander isomorph, und für sie gelten die Aussagen, die in 3.4. bezüglich des abstrakten separablen Hilbertraumes bewiesen wurden.

7.2.5. Der Zusammenhang zwischen der Konvergenz im Quadratmittel und den anderen Arten der Konvergenz von Funktionenfolgen. Ausgehend von der Norm im reellen Raum L_2[1]) definierten wir für quadratisch integrierbare Funktionen folgenden Konvergenzbegriff:

$$f_n \to f,$$

wenn

$$\lim_{n\to\infty} \int [f_n(x) - f(x)]^2 \, d\mu = 0$$

ist. Diese Konvergenz bezüglich der Metrik des Raumes $L_2(X, \mu)$ wird als *Konvergenz im Quadratmittel* bezeichnet. Im folgenden untersuchen wir nun die Beziehungen zwischen dieser Konvergenzart und den anderen Konvergenzarten von Funktionenfolgen. Dabei setzen wir zunächst voraus, daß die Trägermenge X ein endliches Maß hat.

1. *Wenn eine Folge $\{f_n\}$ von Funktionen aus $L_2(X, \mu)$ in der Metrik des Raumes $L_2(X, \mu)$ konvergiert, dann konvergiert sie auch in der Metrik des Raumes $L_1(X, \mu)$.*

Diese Aussage ergibt sich sofort aus der Ungleichung (1), nach der

$$\int |f_n(x) - f(x)| \, d\mu \leqq [\mu(X) \int (f_n(x) - f(x))^2 \, d\mu]^{1/2}$$

ist.

2. *Wenn eine Folge $\{f_n\}$ gleichmäßig auf X konvergiert, dann konvergiert sie auch im Quadratmittel.*

Da für jedes $\varepsilon > 0$ für alle hinreichend großen n

$$|f_n(x) - f(x)| < \varepsilon, \qquad x \in X,$$

ist, gilt

$$\int [f_n(x) - f(x)]^2 \, d\mu < \varepsilon^2 \mu(X),$$

woraus die Behauptung folgt.

3. *Wenn eine Folge $\{f_n\}$ integrierbarer Funktionen im Mittel konvergiert, dann konvergiert sie auf X dem Maß nach.*

Diese Aussage ergibt sich unmittelbar aus der Čebyševschen Ungleichung (vgl. 5.5.4.).

---

[1]) Anm. d. Übers.: Die Betrachtungen gelten entsprechend auch für den komplexen Raum L_2.

Fügt man die dritte Aussage und die des Satzes 8 aus 5.4. zusammen, so folgt:

4. *Wenn eine Folge* $\{f_n\}$ *im Mittel konvergiert, dann kann man aus ihr eine Teilfolge* $\{f_{n_k}\}$ *aussondern, die fast überall konvergiert.*

Diese Tatsache hatten wir auch beim Beweis der Vollständigkeit des Raumes L_1 direkt abgeleitet, ohne Satz 8 aus 5.4. zu benutzen.

Man sieht leicht, daß aus der Konvergenz einer Folge im Mittel (oder sogar im Quadratmittel) im allgemeinen nicht ihre Konvergenz fast überall folgt. So konvergiert die in 5.4.6. konstruierte Folge $\{f_n\}$ im Mittel (und sogar im Quadratmittel) gegen $f \equiv 0$, während sie, wie wir früher sahen, punktweise nirgends gegen 0 konvergiert. Umgekehrt kann auch eine Folge $\{f_n\}$ fast überall (oder sogar überall) konvergieren, ohne daß sie im Mittel konvergiert. Die Funktionenfolge

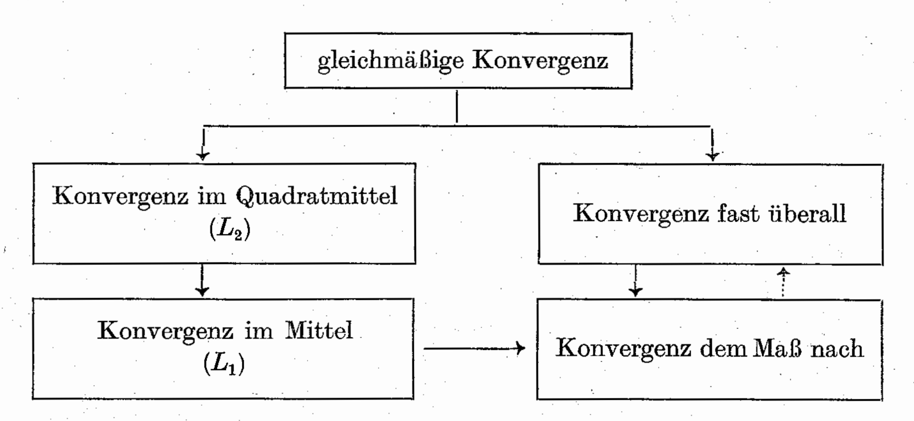
$$f_n(x) = \begin{cases} n & \text{für } x \in \left(0, \dfrac{1}{n}\right], \\ 0 & \text{sonst} \end{cases}$$

konvergiert z. B. gegen 0 in allen Punkten des Intervalls [0, 1], aber nicht im Mittel gegen 0, da für alle n

$$\int_0^1 |f_n(x)|\, dx = 1$$

ist.

Der Zusammenhang zwischen den verschiedenen Konvergenzarten im Fall $\mu(X) < \infty$ kann durch das folgende Schema dargestellt werden:



Dabei bezeichnet der punktierte Pfeil die Möglichkeit der Auswahl einer fast überall konvergenten Teilfolge aus einer Folge, die dem Maß nach konvergiert.

Liegt der Fall $\mu(X) = \infty$ vor (wenn beispielsweise Funktionen auf der ganzen Zahlengeraden mit dem Lebesgueschen Maß betrachtet werden), dann bestehen die

oben dargestellten Beziehungen nicht. So konvergiert z. B. die Folge

$$f_n(x) = \begin{cases} \frac{1}{\sqrt{n}} & \text{für} \quad |x| \leqq n, \\ 0 & \text{für} \quad |x| > n \end{cases}$$

auf der ganzen Zahlengeraden gleichmäßig gegen die Funktion $f \equiv 0$, während sie weder im Mittel noch im Quadratmittel konvergiert. Desweiteren ergibt sich im Fall $\mu(X) = \infty$, wie wir früher bereits gezeigt haben (vgl. 7.2.2.), aus der Konvergenz einer Folge im Quadratmittel (d. h. in L_2) im allgemeinen nicht die Konvergenz dieser Folge im Mittel (d. h. in L_1).

Zum Abschluß bemerken wir noch, daß aus der Konvergenz im Mittel weder im Fall $\mu(X) < \infty$ noch im Fall $\mu(X) = \infty$ die Konvergenz im Quadratmittel folgt.

7.3. Orthogonale Funktionensysteme in L_2. Reihenentwicklungen nach orthogonalen Funktionen

Nach den allgemeinen Sätzen aus 3.4. über unitäre Räume gibt es in L_2 vollständige orthogonale (insbesondere orthonormierte) Funktionensysteme. Man erhält solche Systeme, indem man vollständige Systeme nach dem in 3.4.3. beschriebenen Verfahren orthogonalisiert. Wenn in L_2 ein vollständiges Orthogonalsystem $\{\varphi_n\}$ ausgewählt ist, kann nach den allgemeinen Resultaten aus 3.4. jedes Element $f \in L_2$ in Form einer Reihe

$$f = \sum_{n=1}^{\infty} c_n \varphi_n,$$

der Fourierreihe der Funktion f bezüglich des Systems $\{\varphi_n\}$, dargestellt werden. Dabei sind die Koeffizienten c_n die Fourierkoeffizienten der Funktion f bezüglich des Systems $\{\varphi_n\}$,

$$c_n = \frac{1}{\|\varphi_n\|^2} \int f(x)\, \varphi_n(x)\, d\mu \qquad \left(\|\varphi_n\|^2 = \int \varphi_n^2(x)\, d\mu\right).$$

Im weiteren werden wir nun wichtige Beispiele orthogonaler Funktionensysteme in den Räumen L_2 und die entsprechenden Reihenentwicklungen betrachten.

7.3.1. Das trigonometrische System. Die trigonometrische Fourierreihe. Wir betrachten den Raum $L_2[-\pi, \pi]$ der auf dem Intervall $[-\pi, \pi]$ bezüglich des üblichen Lebesgueschen Maßes quadratisch integrierbaren Funktionen. In diesem Raum bilden die Funktionen

$$1, \cos nx, \sin nx \qquad (n = 1, 2, \ldots) \tag{1}$$

ein vollständiges Orthogonalsystem, das sogenannte *trigonometrische System*. Die Orthogonalität bestätigt man leicht durch direktes Ausrechnen. So ist z. B. für $n \neq m$

$$\int_{-\pi}^{\pi} \cos nx \cos mx \, dx = \frac{1}{2} \int_{-\pi}^{\pi} \left[\cos \frac{n+m}{2} x + \cos \frac{n-m}{2} x \right] dx = 0,$$

usw. Die Vollständigkeit des Systems (1) folgt aus dem Satz von WEIERSTRASS über die Approximation einer beliebigen stetigen periodischen Funktion durch trigonometrische Polynome.[1]) Das System (1) ist nicht normiert. Normiert man die Funktionen von (1), so erhält man das System

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin nx}{\sqrt{\pi}} \qquad (n = 1, 2, \ldots).$$

Ist f eine Funktion aus $L_2[-\pi, \pi]$ und bezeichnen wir ihre Fourierkoeffizienten bezüglich der Funktionen 1, $\cos nx$, $\sin nx$ (um die Beschreibung möglichst einfach zu gestalten) mit $a_0/2$, a_n, b_n, so ergibt sich aus den allgemeinen Formeln

$$\frac{a_0}{2} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \, dx, \quad \text{d. h.} \quad a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \, dx$$

und

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx, \qquad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx \, dx.$$

Die entsprechende Fourierreihe der Funktion $f \in L_2$ hat dann die Form

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

und konvergiert im Quadratmittel gegen f. Ist

$$S_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx)$$

die n-te Partialsumme der Fourierreihe, dann gilt für die mittlere quadratische Abweichung von $S_n(x)$ und $f(x)$

$$\|f(x) - S_n(x)\|^2 = \|f\|^2 - \pi \left(\frac{a_0^2}{2} + \sum_{k=1}^{n} a_k^2 + b_k^2 \right).$$

[1]) In 8.2. beweisen wir den Satz von FEJÉR, der eine Verschärfung des Satzes von WEIERSTRASS ist. Dabei wird auch ein Beweis für die Vollständigkeit des trigonometrischen Systems gegeben, der sich natürlich nicht auf die hier dargestellten Fakten stützt.

Das ist nach 3.4.4. die kleinste mittlere quadratische Abweichung, die bei der Approximation von $f(x)$ durch trigonometrische Polynome

$$T_n(x) = \frac{\alpha_0}{2} + \sum_{k=1}^{n} (\alpha_k \cos kx + \beta_k \sin kx)$$

n-ten Grades erreicht werden kann.

Die Besselsche Ungleichung lautet für das trigonometrische System

$$\frac{a_0^2}{2} + \sum_{n=1}^{\infty} (a_n^2 + b_n^2) \leqq \frac{1}{\pi} \int_{-\pi}^{\pi} f^2(x)\, dx.$$

Da das trigonometrische System vollständig ist, gilt für eine beliebige Funktion aus L_2 sogar die Parsevalsche Gleichung

$$\frac{a_0^2}{2} + \sum_{n=1}^{\infty} (a_n^2 + b_n^2) = \frac{1}{\pi} \int_{-\pi}^{\pi} f^2(x)\, dx,$$

d. h., für eine beliebige Funktion $f \in L_2$ konvergiert die Reihe der Quadrate der Fourierkoeffizienten. Andererseits folgt aus der Konvergenz der Reihe $\sum_{n=1}^{\infty} (a_n^2 + b_n^2)$ für vorgegebene Zahlen a_0, a_n, b_n $(n = 1, 2, \ldots)$ die Konvergenz der Reihe

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

(in L_2) gegen eine Grenzfunktion $f \in L_2$, deren Fourierkoeffizienten gerade die vorgegebenen Zahlen a_0, a_n, b_n $(n = 1, 2, \ldots)$ sind (vgl. Satz 3 aus 3.4.).

Alle bisherigen Aussagen lassen sich leicht auf Funktionen übertragen, die auf einem Intervall $[-l, l]$ beliebiger Länge gegeben sind. Wenn $f(t)$ eine quadratisch integrierbare Funktion auf $[-l, l]$ ist, dann wird $f(t)$ durch die Substitution $t = lx/\pi$, d. h. $x = \pi t/l$, zu einer quadratisch integrierbaren Funktion $f^*(x) = f(lx/\pi)$ auf dem Intervall $[-\pi, \pi]$, und wir erhalten aus den obigen Formeln

$$a_n = \frac{1}{l} \int_{-l}^{l} f(t) \cos \frac{n\pi t}{l}\, dt \qquad (n = 0, 1, \ldots),$$

$$b_n = \frac{1}{l} \int_{-l}^{l} f(t) \sin \frac{n\pi t}{l}\, dt \qquad (n = 1, 2, \ldots)$$

als Fourierkoeffizienten und dazu

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos\left(\frac{n\pi t}{l}\right) + b_n \sin\left(\frac{n\pi t}{l}\right)\right)$$

als Fourierreihe einer Funktion $f \in L_2[-l, l]$.

Bemerkungen

1. Trigonometrische Reihen wurden von dem französischen Mathematiker C. Fourier in seinen Arbeiten über mathematische Physik und dort hauptsächlich bei der Theorie der Wärmeleitung in einem Körper benutzt. Formeln für die Koeffizienten a_n und b_n treten allerdings schon bei Euler auf. Weiterentwickelt wurde die Theorie der trigonometrischen Reihen in den Arbeiten von Riemann, Dirichlet und anderen. Die Begriffe „Fourierreihe", „Fourierkoeffizienten" usw. wurden zuerst nur in Verbindung mit dem trigonometrischen Funktionensystem benutzt. In dem allgemeinen Sinn wie in 3.4. (d. h. bezüglich eines beliebigen orthogonalen Systems in einem beliebigen unitären Raum) wurden sie erst viel später verwendet.

2. Aus der Vollständigkeit des trigonometrischen Systems und den allgemeinen Sätzen aus 3.4. folgt, daß für jede Funktion $f \in L_2$ die zugehörige Fourierreihe

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

im Mittel gegen die gegebene Funktion f konvergiert. Vom Standpunkt konkreter Aufgaben der Analysis ist es jedoch wichtig zu wissen, wann diese Reihe punktweise oder gleichmäßig gegen f konvergiert. Mit diesen Fragen werden wir uns im nächsten Kapitel befassen.

7.3.2. Trigonometrische Systeme auf dem Intervall $[0, \pi]$. Die Funktionen

$$1, \cos x, \cos 2x, \ldots \tag{2}$$

und

$$\sin x, \sin 2x, \ldots \tag{3}$$

bilden in ihrer Gesamtheit ein vollständiges Orthogonalsystem auf dem Intervall $[-\pi, \pi]$. Auf dem Intervall $[0, \pi]$ ist jedes der beiden Systeme (2) und (3) für sich ein vollständiges Orthogonalsystem. Die Orthogonalität dieser Systeme ergibt sich leicht durch direkte Rechnung, ihre Vollständigkeit durch die nachfolgenden Betrachtungen. Zunächst beweisen wir die Vollständigkeit des Systems (2).

Es sei $f(x)$ eine quadratisch integrierbare Funktion auf $[0, \pi]$. Wir setzen diese Funktion durch die Festlegung

$$f(x) = f(-x)$$

auf das Intervall $[-\pi, 0)$ fort und entwickeln die fortgesetzte Funktion bezüglich des Systems

$$1, \cos nx, \sin nx \qquad (n = 1, 2, \ldots)$$

in eine Fourierreihe. Da die fortgesetzte Funktion auf $[-\pi, \pi]$ gerade ist, verschwinden alle Koeffizienten der Sinusglieder in der Fourierreihe. Das folgt sofort aus der Formel für diese Koeffizienten, die für eine gerade Funktion f und $n \geqq 1$

$$\int_{-\pi}^{\pi} f(x) \sin nx \, dx = \int_{-\pi}^{0} f(x) \sin nx \, dx + \int_{0}^{\pi} f(x) \sin nx \, dx$$

$$= -\int_{0}^{\pi} f(x) \sin nx \, dx + \int_{0}^{\pi} f(x) \sin nx \, dx = 0$$

liefert. Also kann die fortgesetzte Funktion auf $[-\pi, \pi]$ (und damit auch die Ausgangsfunktion auf $[0, \pi]$) beliebig gut durch Linearkombinationen der Elemente des Systems (2) im Quadratmittel approximiert werden. Daraus folgt die Vollständigkeit des Systems (2). Die Vollständigkeit des Systems (3) auf dem Intervall $[0, \pi]$ beweisen wir analog. Zuerst setzen wir die Funktion $f(x) \in L_2[0, \pi]$ durch die Festlegung

$$f(x) = -f(-x)$$

auf das Intervall $[-\pi, 0)$ fort und entwickeln nun die fortgesetzte Funktion auf $[-\pi, \pi]$ in eine Fourierreihe. Diese Reihe enthält nur Sinusglieder, da die fortgesetzte Funktion auf $[-\pi, \pi]$ ungerade ist. Daraus ergibt sich wie oben die Vollständigkeit des Systems (3).

7.3.3. Trigonometrische Fourierreihen in komplexer Form. Die trigonometrischen Reihen können in geschlossener Form geschrieben werden, wenn man die Eulerschen Formeln

$$\cos nx = \frac{e^{inx} + e^{-inx}}{2} \quad \text{und} \quad \sin nx = \frac{e^{inx} - e^{-inx}}{2i}$$

benutzt. Setzen wir diese Ausdrücke in eine Fourierreihe ein, so erhalten wir

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx + b_n \sin nx = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \frac{e^{inx} + e^{-inx}}{2} - ib_n \frac{e^{inx} - e^{-inx}}{2} \right)$$

$$= \frac{a_0}{2} + \sum_{n=1}^{\infty} \frac{a_n - ib_n}{2} e^{inx} + \sum_{n=1}^{\infty} \frac{a_n + ib_n}{2} e^{-inx}$$

$$= \sum_{n=-\infty}^{\infty} c_n e^{inx},$$

wobei

$$c_0 = \frac{a_0}{2}$$

und für $n \geqq 1$

$$c_n = \frac{a_n - ib_n}{2}, \qquad c_{-n} = \frac{a_n + ib_n}{2} \tag{4}$$

gesetzt wurde. Der Ausdruck

$$\sum_{n=-\infty}^{\infty} c_n e^{inx}$$

heißt *trigonometrische Fourierreihe in komplexer Form*. Die Koeffizienten c_n dieser Reihe lassen sich aus den Koeffizienten a_n und b_n mit Hilfe der Gleichungen (4) berechnen. Man findet für sie jedoch auch leicht direkte Formeln. Dazu multiplizieren wir die Gleichung

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx} \tag{5}$$

mit e^{-imx} $(m = 0, \pm 1, \pm 2, \ldots)$ und integrieren über $[-\pi, \pi]$. Auf Grund der leicht direkt nachprüfbaren Beziehungen

$$\int_{-\pi}^{\pi} e^{inx} e^{-imx}\, dx = \begin{cases} 0 & \text{für } n \neq m, \\ 2\pi & \text{für } n = m, \end{cases}$$

ergibt sich

$$\int_{-\pi}^{\pi} f(x)\, e^{-imx}\, dx = 2\pi c_m,$$

d. h.

$$c_m = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\, e^{-imx}\, dx \qquad (m = 0, \pm 1, \pm 2, \ldots). \tag{6}$$

Eine Entwicklung der Form (5) ist auch für komplexwertige auf $[-\pi, \pi]$ quadratisch integrierbare Funktionen möglich. Die Funktionen e^{inx} $(n = 0, \pm 1, \pm 2, \ldots)$ bilden also ein vollständiges Orthogonalsystem im Raum $L_2[-\pi, \pi]$ der komplexwertigen Funktionen, deren Absolutbetrag auf $[-\pi, \pi]$ quadratisch integrierbar ist. Die Ausdrücke (6) sind dann die Skalarprodukte der Funktionen f mit den Funktionen e^{imx} im komplexen Raum L_2.

Ersetzen wir die Funktionen e^{inx} durch die Funktionen $e^{i\frac{n\pi}{l}x}$, dann übertragen sich alle Aussagen auf den Raum $L_2[-l, l]$ der komplexwertigen Funktionen, die auf einem Intervall $[-l, l]$ beliebiger Länge definiert sind.

7.3.4. Die Legendreschen Polynome. Die Linearkombinationen der Funktionen

$$1, x, x^2, \dots \tag{7}$$

bilden die Gesamtheit aller Polynome. Daher ist das System (7) im Raum $L_2[a, b]$ der auf dem beliebigen Intervall $[a, b]$ quadratisch integrierbaren Funktionen vollständig.[1] Orthogonalisieren wir das System (7) auf dem Intervall $[-1, 1]$ bezüglich des Skalarproduktes

$$(f, g) = \int_{-1}^{1} f(x)\, g(x)\, dx,$$

so erhalten wir ein vollständiges Orthogonalsystem

$$Q_0(x), Q_1(x), Q_2(x), \dots,$$

wobei $Q_n(x)$ ein Polynom n-ten Grades ist. Wir zeigen nun, daß jedes Polynom $Q_n(x)$ bis auf einen konstanten Faktor mit dem Polynom

$$R_n(x) = \frac{d^n}{dx^n} (x^2 - 1)^n$$

übereinstimmt. Dazu beweisen wir zunächst, daß das System $\{R_n\}$ orthogonal ist. Es sei $n \geqq m$. Dann ist

$$\left.\frac{d^k}{dx^k} (x^2 - 1)^n\right|_{x=-1} = \left.\frac{d^k}{dx^k} (x^2 - 1)^n\right|_{x=1} = 0$$

für alle $k = 0, 1, \dots, n - 1$, und durch fortgesetzte partielle Integration ergibt sich

$$\int_{-1}^{1} R_m(x)\, R_n(x)\, dx = -\int_{-1}^{1} \frac{d^{m+1}}{dx^{m+1}} (x^2 - 1)^m \cdot \frac{d^{n-1}}{dx^{n-1}} (x^2 - 1)^n\, dx = \cdots$$

$$= (-1)^n \int_{-1}^{1} \left[\frac{d^{m+n}}{dx^{m+n}} (x^2 - 1)^m\right] (x^2 - 1)^n\, dx. \tag{8}$$

Für $m < n$ steht unter dem letzten Integral 0, woraus die Orthogonalität des Systems $\{R_n\}$ folgt.

Aus der Tatsache, daß das Polynom R_n den Grad n hat, folgt weiter, daß R_n in dem Unterraum von L_2 liegen muß, der von den ersten $n + 1$ Elementen des Systems (7) erzeugt wird. Damit besitzen die Systeme $\{R_n\}$ und $\{Q_n\}$ folgende gemeinsame Eigenschaften:

1. Sie sind beide orthogonal;
2. das n-te Element des Systems gehört zu dem Unterraum von $L_2[-1, 1]$, der von den Elementen $1, x, \dots, x^{n-1}$ erzeugt wird.

---

[1]) Die Vollständigkeit des Systems der Polynome im Raum $L_2[a, b]$ folgt aus dem Satz von Weierstrass über die gleichmäßige Approximierbarkeit einer beliebigen stetigen Funktion auf $[a, b]$ durch Polynome (vgl. 8.2.2.).

Nach Satz 1 aus Abschnitt 3.4. bestimmen diese beiden Eigenschaften jedes Element des Systems bis auf einen konstanten Faktor, was zu zeigen war.

Um die Normierungsfaktoren für die Polynome $R_n(x)$ zu finden, betrachten wir die Gleichung (8) für den Fall $n = m$:

$$\int_{-1}^{1} R_n^2(x)\,dx = (-1)^n \int_{-1}^{1} \left[\frac{d^{2n}}{dx^{2n}}(x^2-1)^n\right](x^2-1)^n\,dx$$

$$= (2n)!\int_{-1}^{1}(1-x^2)^n\,dx = \frac{(n!)^2\cdot 2^{2n+1}}{2n+1}, \text{ [1]}$$

d. h., die Norm des Polynoms R_n ist gleich $n!\,2^n\sqrt{\frac{2}{2n+1}}$. Damit bilden die Polynome

$$\frac{1}{n!\,2^n}\sqrt{\frac{2n+1}{2}}\,R_n(x)$$

ein vollständiges Orthonormalsystem im Raum $L_2[-1, 1]$.

Gewöhnlich werden jedoch nicht diese normierten Polynome betrachtet, sondern die durch die Formel

$$P_n(x) = \frac{1}{n!\,2^n}\frac{d^n}{dx^n}(x^2-1)^n$$

definierten Polynome. Die Polynome $P_n(x)$ werden *Legendresche Polynome*, die Formel selbst *Formel von Rodrigues* genannt. Aus den obigen Betrachtungen folgt

$$\int_{-1}^{1} P_n(x)\,P_m(x)\,dx = \begin{cases} 0 & \text{für } n \neq m, \\ \dfrac{2}{2n+1} & \text{für } n = m. \end{cases}$$

Die expliziten Darstellungen für die ersten fünf Legendreschen Polynome lauten

$$P_0(x) = 1,$$

$$P_1(x) = x,$$

$$P_2(x) = \frac{3}{2}x^2 - \frac{1}{2},$$

$$P_3(x) = \frac{5}{2}x^3 - \frac{3}{2}x,$$

$$P_4(x) = \frac{35}{8}x^4 - \frac{15}{4}x^2 + \frac{3}{8}.$$

[1]) Das letzte Integral kann man elementar mit Hilfe einer Rekursionsformel berechnen oder auf die Betafunktion zurückführen.

Die Entwicklung einer Funktion $f \in L_2[-1, 1]$ nach Legendreschen Polynomen hat die Gestalt

$$f(x) = \sum_{n=0}^{\infty} c_n P_n(x)$$

mit

$$c_n = \frac{2n+1}{2} \int_{-1}^{1} f(x) P_n(x)\, dx.$$

7.3.5. Orthogonale Funktionensysteme über Mengenprodukten. Mehrfache Fourierreihen. Auf den Mengen X' und X'' seien die Maße μ' und μ'' definiert. L_2' und L_2'' seien die entsprechenden Räume der quadratisch integrierbaren Funktionen über diesen Mengen. Über dem Mengenprodukt

$$X = X' \times X''$$

mit dem Maß

$$\mu = \mu' \times \mu''$$

betrachten wir den entsprechenden Raum L_2 der quadratisch integrierbaren Funktionen, die wir als Funktionen von zwei Variablen schreiben werden.

Satz 1. *Sind $\{\varphi_m\}$ und $\{\psi_n\}$ vollständige Orthonormalsysteme in L_2' bzw. L_2'', dann ist das System aller Produkte*

$$f_{mn}(x, y) = \varphi_m(x)\, \psi_n(y)$$

ein vollständiges Orthonormalsystem in L_2.

Beweis. Nach dem Satz von Fubini ist

$$\int_X f_{mn}^2(x, y)\, d\mu = \int_{X'} \varphi_m^2(x) \left(\int_{X''} \psi_n^2(y)\, d\mu'' \right) d\mu' = 1.$$

Wenn $m \neq m_1$ ist, dann ergibt sich nach demselben Satz

$$\int_X f_{mn}^2(x, y)\, f_{m_1 n_1}(x, y)\, d\mu = \int_{X''} \psi_n(y)\, \psi_{n_1}(y) \left(\int_{X'} \varphi_m(x)\, \varphi_{m_1}(x)\, d\mu' \right) d\mu'' = 0,$$

da die Funktion

$$f_{mn}(x, y)\, f_{m_1 n_1}(x, y)$$

auf $X = X' \times X''$ integrierbar ist.

Wenn $m = m_1$ und $n \neq n_1$ ist, dann gilt ebenfalls

$$\int_X f_{mn}(x, y)\, f_{m_1 n_1}(x, y)\, d\mu = \int_{X'} \varphi_m^2(x) \left(\int_{X''} \psi_n(y)\, \psi_{n_1}(y)\, d\mu'' \right) d\mu' = 0.$$

Wir beweisen nun die Vollständigkeit des Systems $\{f_{mn}\}$. Dazu nehmen wir an, daß es in L_2 eine Funktion f gibt, die zu allen Funktionen f_{mn} orthogonal ist. Setzen wir

$$F_m(y) = \int_{X'} f(x, y)\, \varphi_m(x)\, d\mu',$$

so sieht man leicht, daß die Funktion $F_m(y)$ quadratisch integrierbar ist. Damit ist $F_m(y)\, \psi_n(y)$ für beliebiges n integrierbar, und wir erhalten mit Hilfe des Satzes von FUBINI

$$\int_{X''} F_m(y)\, \psi_n(y)\, d\mu'' = \int_X f(x, y)\, f_{mn}(x, y)\, d\mu = 0.$$

Aus der Vollständigkeit des Systems $\{\psi_n\}$ folgt nun für fast alle y und jedes m

$$F_m(y) = 0,$$

d. h.

$$\int_{X'} f(x, y)\, \varphi_m(x)\, d\mu' = 0.$$

Da auch das System $\{\varphi_m\}$ vollständig ist, ergibt sich hieraus, daß für fast alle y die Menge derjenigen x, für die

$$f(x, y) \neq 0$$

ist, das Maß Null hat. Das bedeutet nach dem Satz von FUBINI, daß die Funktion $f(x, y)$ fast überall gleich Null auf X ist, was zu zeigen war.

Wenden wir den Satz auf den Fall der quadratisch integrierbaren Funktionen zweier reeller Variabler

$$f(x, y) \qquad (-\pi \leqq x, y \leqq \pi)$$

an, dann folgt: Die paarweise gebildeten Produkte der Elemente der Systeme

$$1, \cos mx, \sin mx \qquad (m = 1, 2, \ldots)$$

und

$$1, \cos ny, \sin ny \qquad (n = 1, 2, \ldots),$$

d. h. die Funktionen

$$1, \cos mx, \sin mx, \cos ny, \sin ny,$$

$$\cos mx \cos ny, \cos mx \sin ny,$$

$$\sin mx \cos ny, \sin mx \sin ny$$

bilden ein vollständiges Orthogonalsystem im Raum $L_2([-\pi, \pi] \times [-\pi, \pi])$. Die entsprechende Fourierreihe sieht für diese Entwicklungsfunktionen ein wenig umständlich aus. Wesentlich übersichtlicher wird die Darstellung, wenn wir Exponential-

funktionen benutzen. Dann können die Produkte der Elemente der Orthogonalsysteme in der Form

$$e^{imx}e^{iny} = e^{i(mx+ny)} \qquad (n, m = 0, \pm 1, \pm 2, \ldots)$$

und die entsprechende Fourierreihe in der Form

$$f(x, y) = \sum_{m,n=-\infty}^{\infty} c_{mn} e^{i(mx+ny)}$$

geschrieben werden, wobei

$$c_{mn} = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f(x, y)\, e^{-i(mx+ny)}\, dx\, dy$$

ist.

Die Legendreschen Polynome liefern im Raum $L_2(Q)$ der auf dem Quadrat

$$Q = \{(x, y)\colon -1 \leqq x, y \leqq 1\}$$

quadratisch integrierbaren Funktionen ein vollständiges Orthonormalsystem, das aus den Polynomen

$$Q_{mn}(x, y) = \frac{\sqrt{(2m+1)(2n+1)}}{m!\, n!\, 2^{m+n+1}} \cdot \frac{d^m}{dx^m} (x^2 - 1)^m \cdot \frac{d^n}{dx^n} (y^2 - 1)^n$$

besteht.

Alles oben Gesagte ist offensichtlich auf Funktionen mehrerer Variabler übertragbar. Insbesondere ergibt sich für die trigonometrische Fourierreihe einer Funktion von k reellen Variablen die Form

$$f(x_1, \ldots, x_k) = \sum_{n_1,\ldots,n_k=-\infty}^{\infty} c_{n_1\ldots n_k} e^{i(n_1x_1+\cdots+n_kx_k)}$$

mit

$$c_{n_1\ldots n_k} = \frac{1}{(2\pi)^k} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} f(x_1, \ldots, x_k)\, e^{-i(n_1x_1+\cdots+n_kx_k)}\, dx_1 \cdots dx_k.$$

7.3.6. Polynome, die bezüglich eines vorgegebenen Gewichts orthogonal sind. Durch Orthogonalisierung der Funktionen

$$1, x, x^2, \ldots, x^n, \ldots \tag{9}$$

bezüglich des Skalarprodukts

$$\int_{-1}^{1} f(x)\, g(x)\, dx$$

mit dem gewöhnlichen Lebesgueschen Maß dx auf $[-1, 1]$ kommen wir zu den Legendreschen Polynomen. Wenn auf diesem Intervall ein anderes Maß μ gegeben ist und die Funktionen (9) im entsprechenden Raum $L_2([-1, 1], \mu)$ mit dem Skalarprodukt $\int_{-1}^{1} f(x)\, g(x)\, d\mu$ linear unabhängig sind, dann führt der Orthogonalisierungsprozeß, ausgehend vom System (9), zu einem System von Polynomen $\{Q_n\}$. Dieses System hängt von der Wahl des Maßes μ ab. Setzen wir voraus, daß das Maß μ für alle Lebesgue-meßbaren Untermengen von $[-1, 1]$ durch die Formel

$$\mu(E) = \int_E g(x)\, dx \tag{10}$$

definiert ist ($g(x)$ eine feste nichtnegative integrierbare Funktion), dann lautet die Orthogonalitätsbedingung für das System $\{Q_n\}$,

$$(Q_m, Q_n) = \begin{cases} 1 & \text{für} \quad m = n, \\ 0 & \text{für} \quad m \neq n, \end{cases}$$

in diesem Fall

$$\int_{-1}^{1} Q_m(x)\, Q_n(x)\, g(x)\, dx = \begin{cases} 1 & \text{für} \quad m = n, \\ 0 & \text{für} \quad m \neq n. \end{cases} \tag{11}$$

Die Funktion $g(x)$, die das Maß (10) definiert, wird als *Gewicht* oder *Gewichtsfunktion* bezeichnet. Von den Polynome Q_n, die der Bedingung (11) genügen, sagt man, daß sie orthogonal bezüglich des Gewichts g sind. Die Verwendung verschiedener Gewichte führt zu verschiedenen Systemen von orthogonalen Polynomen. Setzen wir speziell

$$g(x) = \frac{1}{\sqrt{1 - x^2}},$$

so erhalten wir Polynome, die bis auf einen konstanten Faktor mit den sogenannten *Čebyševschen Polynomen*

$$T_n(x) = \cos\,(n \arccos x) \qquad (n = 1, 2, \ldots)$$

übereinstimmen. Diese Polynome spielen eine wichtige Rolle in verschiedenen Approximationsaufgaben. Ihre Orthogonalität bezüglich des Gewichts $1/\sqrt{1 - x^2}$ ergibt sich sofort, wenn wir

$$x = \cos\theta, \qquad dx = -\sin\theta\, d\theta, \qquad \sqrt{1 - x^2} = \sin\theta$$

substituieren. Wir erhalten so:

$$\int_{-1}^{1} \frac{T_m(x)\, T_n(x)}{\sqrt{1 - x^2}}\, dx = \int_0^{\pi} \cos m\theta \cos n\theta\, d\theta = \begin{cases} \dfrac{\pi}{2} & \text{für} \quad m = n, \\ 0 & \text{für} \quad m \neq n. \end{cases}$$

7.3.7. Eine orthogonale Basis im Raum $L_2(-\infty, \infty)$. Die Hermiteschen Funktionen. Oben haben wir Orthogonalsysteme auf einem Intervall, d. h. auf einer Menge endlichen Maßes betrachtet. Jetzt betrachten wir den Fall einer Menge unendlichen Maßes, wir versuchen Orthogonalsysteme im Raum $L_2(-\infty, \infty)$ der auf der ganzen Zahlengeraden quadratisch integrierbaren Funktionen zu finden. Ein solches Funktionensystem kann weder aus Polynomen noch aus trigonometrischen Funktionen noch aus irgendwelchen anderen Funktionen konstruiert werden, die nicht zu diesem Raum gehören. Als „Material" für die Konstruktion einer Basis in $L_2(-\infty, \infty)$ können nur Funktionen verwendet werden, die für $|x| \to \infty$ genügend schnell gegen 0 streben. So kann man insbesondere, ausgehend von den in L_2 linear unabhängigen Funktionen

$$x^n e^{-x^2/2} \qquad (n = 0, 1, 2, \ldots),$$

eine Basis in $L_2(-\infty, \infty)$ gewinnen. Durch Orthogonalisierung erhalten wir ein System von Funktionen der Form

$$\varphi_n(x) = H_n(x)\, e^{-x^2/2} \qquad (n = 0, 1, 2, \ldots), \tag{12}$$

wobei $H_n(x)$ ein Polynom n-ter Ordnung ist. Alle diese Funktionen gehören offensichtlich zu $L_2(-\infty, \infty)$ und bilden, wie in 8.4.3. gezeigt wird, ein vollständiges System. Die Polynome $H_n(x)$ heißen *Hermitesche Polynome*, die Funktionen φ_n *Hermitesche Funktionen*. Die Hermiteschen Polynome stimmen bis auf eine multiplikative Konstante mit den Polynomen

$$H_n^*(x) = (-1)^n e^{x^2} \frac{d^n e^{-x^2}}{dx^n}$$

überein, denn die H_n^* sind offensichtlich Polynome n-ten Grades und erfüllen die Relation

$$\int_{-\infty}^{\infty} H_m^*(x)\, H_n^*(x)\, e^{-x^2}\, dx = 0 \qquad (m \neq n),$$

wie man leicht durch partielle Integration zeigt. Aus 3.4., Satz 1, folgt nun, daß die Funktionen $H_n^*(x)\, e^{-x^2/2}$ und $H_n(x)\, e^{-x^2/2}$ der aus $x^n e^{-x^2/2}$ $(n = 0, 1, 2, \ldots)$ erzeugbaren Orthogonalsysteme bis auf einen konstanten Faktor übereinstimmen.

Das obige Resultat kann auch anders interpretiert werden. Betrachten wir auf der Zahlengeraden das *endliche* Maß μ mit der Dichte e^{-x^2}, d. h.

$$d\mu = e^{-x^2}\, dx,$$

und den Raum $L_2\big((-\infty, \infty), \mu\big)$ mit dem Skalarprodukt

$$(f, g) = \int_{-\infty}^{\infty} f(x)\, g(x)\, e^{-x^2}\, dx,$$

dann lautet das Resultat: Die Hermiteschen Polynome bilden ein vollständiges Orthogonalsystem im Raum $L_2\big((-\infty, \infty), \mu\big)$.

Zum Schluß betrachten wir noch den Raum $L_2(0, \infty)$ der auf der Halbgeraden quadratisch integrierbaren Funktionen. Ausgehend vom System der Funktionen

$$x^n e^{-x} \qquad (n = 0, 1, 2, \ldots)$$

erhalten wir durch Orthogonalisierung das Funktionensystem

$$L_n(x)\, e^{-x}$$

der sogenannte *Laguerreschen Funktionen*. Die Funktionen $L_n(x)$ sind Polynome n-ten Grades und heißen *Laguerresche Polynome*. Diese Polynome kann man ebenfalls als orthogonale Basis im Raum $L_2\big((0, \infty), \mu\big)$ bezüglich des Maßes μ mit der Dichte e^{-x}, d. h.

$$d\mu = e^{-x}\, dx,$$

auffassen. Die Vollständigkeit des Systems der Laguerreschen Funktionen im Raum $L_2(0, \infty)$ wird in 8.4.3. gezeigt.

7.3.8. Polynome, die bezüglich eines diskreten Gewichts orthogonal sind. In $n + 1$ verschiedenen Punkten $x_0, x_1, \ldots, x_n$ der reellen Achse seien die positiven Zahlen $p_0, p_1, \ldots, p_n$ als „Gewichte" gegeben. Das Maß μ sei durch die Formel

$$\mu(E) = \sum_{x_k \in E} p_k$$

für eine beliebige Untermenge E der reellen Achse definiert, d. h., $\mu(E)$ ist gleich der Summe der Gewichte der Punkte, die in der Menge E liegen. Meßbar bezüglich dieses ausgearteten Maßes sind beliebige Mengen und Funktionen auf der Geraden, wobei eine beliebige Menge, die keinen der Punkte x_k $(k = 0, 1, \ldots, n)$ enthält, das Maß 0 besitzt. Das entsprechende Integral einer Funktion f über die reelle Achse ist gleich

$$\int_{-\infty}^{\infty} f(x)\, d\mu = \sum_{k=0}^{n} p_k f(x_k),$$

das Skalarprodukt in $L_2\big((-\infty, \infty), \mu\big)$ gleich

$$(f, g) = \sum_{k=0}^{n} p_k f(x_k)\, g(x_k).$$

Zwei Funktionen f und g sind genau dann äquivalent bezüglich dieses Maßes μ, wenn

$$f(x_k) = g(x_k)$$

in allen Punkten x_k $(k = 0, 1, \ldots, n)$ ist. Die Aufgabe, eine gegebene Funktion f möglichst gut durch eine Linearkombination

$$c_0\varphi_0 + c_1\varphi_1 + \cdots + c_m\varphi_m$$

von Funktionen $\varphi_0, \ldots, \varphi_m$ im Sinne von L_2 zu approximieren, führt für dieses Maß auf das Problem, den Ausdruck

$$\sum_{k=0}^{n} p_k\{f(x_k) - \sum_{i=1}^{m} c_i \varphi_i(x_k)\}^2$$

zu einem Minimum zu machen, d. h., sie führt zum Problem der „Interpolation nach der Methode der kleinsten Quadrate".

Im Zusammenhang mit dem Problem der Approximation einer Funktion durch Polynome nach der Methode der kleinsten Quadrate entwickelte P. L. ČEBYŠEV eine Theorie der orthogonalen Polynome. Um die diesbezüglichen Resultate ČEBYŠEVS darzustellen, bemerken wir zuerst, daß das System

$$1, x, x^2, \ldots, x^n \tag{13}$$

für das oben beschriebene ausgeartete Maß μ linear unabhängig im Raum L_2 ist. Das folgt daraus, daß die Gramsche Determinante $\det\{(x^r, x^s)\}_{r,s=0,\ldots,n}$ des Systems (13) für das Skalarprodukt

$$(x^r, x^s) = \sum_{k=0}^{n} p_k x_k^{r+s}$$

von 0 verschieden ist, wie die folgende Rechnung zeigt:

$$\begin{vmatrix} \sum p_k & \sum p_k x_k & \sum p_k x_k^2 & \cdots & \sum p_k x_k^n \\ \sum p_k x_k & \sum p_k x_k^2 & \sum p_k x_k^3 & \cdots & \sum p_k x_k^{n+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \sum p_k x_k^n & \sum p_k x_k^{n+1} & \sum p_k x_k^{n+2} & \cdots & \sum p_k x_k^{2n} \end{vmatrix}$$

$$= \begin{vmatrix} \sqrt{p_0} & \sqrt{p_1} & \cdots & \sqrt{p_n} \\ \sqrt{p_0}x_0 & \sqrt{p_1}x_1 & \cdots & \sqrt{p_n}x_n \\ \cdots & \cdots & \cdots & \cdots \\ \sqrt{p_0}x_0^n & \sqrt{p_1}x_1^n & \cdots & \sqrt{p_n}x_n^n \end{vmatrix}^2 = p_0 p_1 \cdots p_n \begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_0 & x_1 & \cdots & x_n \\ \cdots & \cdots & \cdots & \cdots \\ x_0^n & x_1^n & \cdots & x_n^n \end{vmatrix}^2 \neq 0.$$

Die Funktionen x^r für $r > n$ sind von den Funktionen des Systems (13) linear abhängig, weil $L_2\big((-\infty, \infty), \mu\big)$ in unserem Fall die Dimension $n + 1$ besitzt. Damit führt der Orthogonalisierungsprozeß, ausgehend vom System (13), zu einem System von Polynomen

$$P_0(x), P_1(x), \ldots, P_n(x),$$

die orthonormal in dem Sinne sind, daß

$$\sum_{k=0}^{n} p_k P_r(x_k)\, P_s(x_k) = \delta_{rs}$$

ist, und für jede Funktion f eine endliche Reihenentwicklung

$$f \sim \sum_{r=0}^{n} c_r P_r(x), \qquad c_r = \sum_{k=0}^{n} p_k P_r(x_k) f(x_k),$$

liefern. „Konvergenz" dieser Reihenentwicklung in dem betrachteten Raum L_2 bedeutet jetzt, daß in den Punkten x_k

$$f(x_k) = \sum_{r=0}^{n} c_r P_r(x_k) \qquad (k = 0, 1, \ldots, n)$$

ist, d. h., die vollständige Summe $\sum_{r=0}^{n} c_r P_r(x)$ ist einfach das Lagrangesche Interpolationspolynom n-ten Grades für die Funktion $f(x)$.

Die nichtvollständigen Summen

$$Q_m(x) = \sum_{r=0}^{m} c_r P_r(x) \qquad (m < n)$$

sind Polynome m-ten Grades, die $f(x)$ in dem Sinne approximieren, daß der Ausdruck

$$\sum_{k=0}^{n} p_k \{f(x_k) - Q_m(x_k)\}^2$$

für $Q_m(x)$ kleiner ist als für jedes andere Polynom m-ten Grades.

7.3.9. Die Funktionensysteme von Haar, Rademacher und Walsh. Haar hat folgendes Beispiel eines vollständigen Funktionensystems in $L_2[0, 1]$ konstruiert. Das System besteht aus der Funktion

$$\varphi_0 = 1$$

und den Serien von Funktionen

$$\begin{array}{l} \varphi_{01}, \\ \varphi_{11}, \varphi_{12}, \\ \varphi_{21}, \varphi_{22}, \varphi_{23}, \varphi_{24}, \\ \cdots\cdots\cdots\cdots \\ \varphi_{n1}, \varphi_{n2}, \ldots, \varphi_{n2^n} \end{array}$$

(die n-te Serie enthält 2^n Funktionen) mit

$$\varphi_{01} = \begin{cases} +1 & \text{für } 0 < x < \frac{1}{2}, \\ -1 & \text{für } \frac{1}{2} < x < 1, \end{cases}$$

$$\varphi_{11} = \begin{cases} \sqrt{2} & \text{für } 0 < x < \frac{1}{4}, \\ -\sqrt{2} & \text{für } \frac{1}{4} < x < \frac{1}{2}, \\ 0 & \text{für } \frac{1}{2} < x < 1, \end{cases} \qquad \varphi_{12} = \begin{cases} 0 & \text{für } 0 < x < \frac{1}{2}, \\ \sqrt{2} & \text{für } \frac{1}{2} < x < \frac{3}{4}, \\ -\sqrt{2} & \text{für } \frac{3}{4} < x < 1 \end{cases}$$

und allgemein

$$\varphi_{ni} = \begin{cases} 2^{n/2} & \text{für } \dfrac{i-1}{2^n} < x < \dfrac{i-1}{2^n} + \dfrac{1}{2^{n+1}}, \\ -2^{n/2} & \text{für } \dfrac{i-1}{2^n} + \dfrac{1}{2^{n+1}} < x < \dfrac{i}{2^n}, \\ 0 & \text{für } x \notin \left[\dfrac{i-1}{2^n}, \dfrac{i}{2^n}\right] \end{cases}$$

$(n = 0, 1, 2, \ldots;\ i = 1, 2, \ldots, 2^n)$.

Man sieht leicht, daß dieses System orthonormiert ist. Wir beweisen nun seine Vollständigkeit. Dazu zerlegen wir das Intervall $[0, 1]$ in 2^{n+1} gleich Intervalle $\varDelta_i$ und betrachten die Menge M_{n+1} aller Funktionen, die auf jedem der Intervalle $\varDelta_i$ jeweils konstant sind. Offensichtlich ist M_{n+1} ein linearer Unterraum der Dimension 2^{n+1} von $L_2[0, 1]$. Außerdem liegen alle Funktionen des Systems bis zur n-ten Serie in diesem Unterraum. Da diese Funktionen auf Grund der Orthonormiertheit linear unabhängig sind und ihre Anzahl gleich

$$1 + \sum_{k=0}^{n} 2^k = 2^{n+1}$$

ist, bilden die Funktionen φ_0 und φ_{ki} der Serien $k = 0, 1, \ldots, n$ ein vollständiges linear unabhängiges System in M_{n+1}. Zusammen mit der Tatsache, daß sich jede stetige Funktion auf $[0, 1]$ beliebig gut durch Funktionen aus M_{n+1} (n hinreichend groß) approximieren läßt, folgt daraus die Vollständigkeit des Haarschen Systems.

Wir betrachten nun noch ein anderes Beispiel eines orthonormierten Systems in $L_2[0, 1]$, das von Rademacher angegeben wurde. Wir setzen

$$\varphi_m(x) = (-1)^{[2^m x]},$$

d. h., die Funktion $\varphi_m(x)$ ergibt sich, indem man das Intervall $[0, 1]$ in 2^m gleiche Teile $\varDelta_i$ teilt und die Funktionswerte von φ_m auf den Intervallen $\varDelta_i$ $(i = 1, \ldots, 2^m)$ abwechselnd $+1$ und -1 ansetzt. Das System

$$\varphi_0, \varphi_1, \ldots, \varphi_n, \ldots \tag{14}$$

ist offensichtlich orthonormiert, aber nicht vollständig, da z. B. die Funktion

$$\varphi_{12} = \varphi_1\varphi_2 = \begin{cases} 1 & \text{für } 0 < x < \dfrac{1}{4} \text{ oder } \dfrac{3}{4} < x < 1, \\ -1 & \text{für } \dfrac{1}{4} < x < \dfrac{3}{4} \end{cases}$$

zu allen Funktionen des Systems (14) orthogonal ist. Man kann jedoch dieses System zu einem vollständigen orthogonalen System erweitern, indem man zu den Funktionen (14) noch alle Funktionen der Form

$$\varphi_{m_1 m_2 \cdots m_k} = \varphi_{m_1}\varphi_{m_2} \cdots \varphi_{m_k} \qquad (0 < m_1 < m_2 < \cdots < m_k) \tag{15}$$

hinzunimmt. Das auf diese Weise erweiterte System heißt *Walshsches System*. Es ist offensichtlich orthonormiert und außerdem vollständig. Das erhält man leicht durch analoge Überlegungen wie beim Beweis der entsprechenden Eigenschaft für das Haarsche System.

8. Trigonometrische Reihen. Die Fouriertransformation

8.1. Konvergenzbedingungen für die Fourierreihe

8.1.1. Hinreichende Bedingungen für die Konvergenz der Fourierreihe in einem Punkt. Wir betrachten wieder den Raum $L_2[-\pi, \pi]$ der quadratisch integrierbaren Funktion auf dem Intervall $[-\pi, \pi]$. Das ist, wie in Kapitel 7 gezeigt wurde, ein vollständiger unendlichdimensionaler unitärer Raum, d. h. ein Hilbertraum. Die Funktionen

$$1, \cos nx, \sin nx \qquad (n = 1, 2, \ldots) \tag{1}$$

bilden in diesem Raum ein vollständiges orthogonales System. Für jede Funktion $f \in L_2[-\pi, \pi]$ konvergiert daher die Fourierreihe

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \tag{2}$$

mit den Koeffizienten

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx, \qquad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx \, dx \tag{3}$$

im quadratischen Mittel, d. h. in der Norm des Raumes $L_2[-\pi, \pi]$ gegen f. Jedoch für die Anwendung von Fourierreihen auf Probleme der mathematischen Physik und andere Fragen sind oft weitergehende Konvergenzbedingungen von Bedeutung. Und zwar interessiert man sich für Bedingungen, die nicht nur die Konvergenz im Mittel, sondern auch die Konvergenz in einem gegebenen Punkt, in allen Punkten oder sogar die gleichmäßige Konvergenz garantieren. Wir werden zunächst hinreichende Bedingungen für die Konvergenz der Fourierreihe in einem gegebenen Punkt aufstellen. Dazu machen wir einige vorbereitende Bemerkungen.

Anstelle von Funktionen auf dem Intervall $[-\pi, \pi]$ können wir auch periodische Funktionen mit der Periode 2π auf der ganzen reellen Achse betrachten, denn jede Funktion auf einem Intervall kann man periodisch fortsetzen. Ferner kann man mit Hilfe der Formeln (3) sogar für jede *integrierbare* Funktion (und nicht nur für jede quadratisch integrierbare Funktion) Fourierkoeffizienten definieren,[1]) denn die trigonometrischen Funktionen (1) sind beschränkt, und daher besitzen die Formeln (3) auch in diesem Fall einen Sinn. Somit entspricht jeder Funktion $f \in L_1[-\pi, \pi]$ eine

[1]) Dabei können wir natürlich für eine beliebige integrierbare Funktion keinerlei Aussagen über die Konvergenz der Reihe (2) machen.

Fourierreihe,

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx).$$

Wir wollen jetzt die Frage untersuchen, ob diese Reihe in einem gegebenen Punkt x gegen den entsprechenden Funktionswert $f(x)$ konvergiert. Dazu setzen wir

$$S_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx). \tag{4}$$

Zunächst formen wir $S_n(x)$ um, indem wir die Koeffizienten a_k und b_k in (4) durch ihre Integralausdrücke (3) ersetzen. Bezeichnen wir dabei die Integrationsveränderliche mit t, so erhalten wir

$$S_n(x) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \left\{\frac{1}{2} + \sum_{k=1}^{n} \cos kx \cos kt + \sin kx \sin kt\right\} dt$$

$$= \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \left\{\frac{1}{2} + \sum_{k=1}^{n} \cos k(t-x)\right\} dt.$$

Unter Verwendung der wohlbekannten Formel[1])

$$\frac{1}{2} + \cos u + \cos 2u + \cdots + \cos nu = \frac{\sin \frac{2n+1}{2} u}{2 \sin \frac{u}{2}} \tag{5}$$

erhalten wir daraus

$$S_n(x) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \frac{\sin \frac{2n+1}{2}(t-x)}{2 \sin \frac{t-x}{2}} dt. \tag{6}$$

---

[1]) Diese Formel ergibt sich etwa durch Summierung der Gleichungen

$$\sin \frac{u}{2} = \frac{1}{2} \cdot 2 \sin \frac{u}{2},$$

$$\sin \frac{3u}{2} - \sin \frac{u}{2} = \cos u \cdot 2 \sin \frac{u}{2},$$

$$\cdots\cdots\cdots\cdots$$

$$\sin \frac{2n+1}{2} u - \sin \frac{2n-1}{2} u = \cos nu \cdot 2 \sin \frac{u}{2}.$$

Diese Darstellung von $S_n(x)$ und verschiedene Modifikationen davon nennt man *Dirichletsches Integral.*

Da unter dem Integral in (6) eine periodische Funktion mit der Periode 2π steht, besitzt das Integral über ein beliebiges Intervall der Länge 2π stets ein und denselben Wert. Setzen wir also $t - x = z$, so können wir auch bei Integration über z die Integrationsgrenzen $-\pi$ und π beibehalten, und wir erhalten

$$S_n(x) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x+z) \frac{\sin \frac{2n+1}{2} z}{2 \sin \frac{z}{2}} \, dz.$$

Die Funktion

$$D_n(z) = \frac{1}{2\pi} \frac{\sin \frac{2n+1}{2} z}{\sin \frac{z}{2}}$$

nennt man den *Dirichletschen Kern.* Aus Gleichung (5) ist unmittelbar zu ersehen, daß für jedes n

$$\int_{-\pi}^{\pi} D_n(z) \, dz = 1$$

ist. Unter Verwendung dieser Gleichung können wir die Differenz $S_n(x) - f(x)$ auch in der Form

$$S_n(x) - f(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} [f(x+z) - f(x)] \frac{\sin \frac{2n+1}{2} z}{\sin \frac{z}{2}} \, dz \tag{7}$$

schreiben. Somit haben wir die Frage der Konvergenz von $S_n(x)$ gegen $f(x)$ auf die Frage zurückgeführt, wann das Integral (7) gegen Null strebt. Die Untersuchung dieses Integrals stützt sich auf das folgende Lemma.

Lemma 1. *Ist φ auf dem Intervall $[a, b]$ integrierbar, so ist*

$$\lim_{p \to \infty} \int_a^b \varphi(x) \sin px \, dx = 0.$$

Beweis. Ist φ eine stetig differenzierbare Funktion, so erhalten wir nach partieller Integration für $p \to \infty$

$$\int_a^b \varphi(x) \sin px \, dx = -\varphi(x) \frac{\cos px}{p}\Big|_a^b + \int_a^b \varphi'(x) \frac{\cos px}{p} \, dx \to 0. \tag{8}$$

Jetzt sei φ eine beliebige auf $[a, b]$ integrierbare Funktion. Da die stetig differenzierbaren Funktionen in $L_1[a, b]$ dicht liegen, kann man zu jedem $\varepsilon > 0$ eine stetig differenzierbare Funktion φ_ε mit

$$\int_a^b |\varphi(x) - \varphi_\varepsilon(x)| \, dx < \frac{\varepsilon}{2} \tag{9}$$

finden. Damit ergibt sich die Abschätzung

$$\begin{aligned}\left|\int_a^b \varphi(x) \sin px \, dx\right| &\leqq \left|\int_a^b [\varphi(x) - \varphi_\varepsilon(x)] \sin px \, dx\right| \\ &\quad + \left|\int_a^b \varphi_\varepsilon(x) \sin px \, dx\right| \\ &\leqq \frac{\varepsilon}{2} + \left|\int_a^b \varphi_\varepsilon(x) \sin px \, dx\right|.\end{aligned}$$

Der zweite Summand strebt auf Grund von (8) für $p \to \infty$ gegen Null. Damit ist das Lemma bewiesen.

Jetzt können wir leicht die folgende hinreichende Bedingung für die Konvergenz der Fourierreihe beweisen.

Satz 1. *Ist f eine integrierbare Funktion und existiert das Integral*

$$\int_{-\delta}^{\delta} \left|\frac{f(x+t) - f(x)}{t}\right| dt \tag{10}$$

bei festem x für ein $\delta > 0$, so konvergieren die Partialsummen S_n der Fourierreihe von f im Punkt x gegen $f(x)$.

Beweis. Wir schreiben das Integral (7) in der Form

$$\frac{1}{\pi} \int_{-\pi}^{\pi} \frac{f(x+z) - f(x)}{z} \frac{z}{2 \sin \frac{z}{2}} \sin \frac{2n+1}{2} z \, dz. \tag{11}$$

Ist die Funktion

$$\frac{f(x+z)-f(x)}{z}$$

in den Grenzen von $-\delta$ bis δ (bezüglich z) integrierbar, so ist sie auch im ganzen Intervall $[-\pi, \pi]$ integrierbar (wegen $f \in L_1[-\pi, \pi]$). Dann ist aber auch die Funktion

$$\frac{f(x+z)-f(x)}{z} \cdot \frac{z}{2\sin\frac{z}{2}}$$

integrierbar. Daher kann man Lemma 1 auf das Integral (11) anwenden, und folglich strebt dieses Integral für $n \to \infty$ gegen Null. Damit ist der Satz bewiesen.

Bemerkungen

1. Die Konvergenz des Integrals (10) nennt man *Dinische Bedingung*. Diese ist insbesondere dann erfüllt, wenn die Funktion f im gegebenen Punkt x stetig ist und eine endliche Ableitung besitzt oder wenigstens eine endliche rechtsseitige und linksseitige Ableitung.

Die beim Beweis von Satz 1 angestellten Überlegungen bleiben gültig, wenn man an Stelle der Dinischen Bedingung die Konvergenz der Integrale

$$\int_{-\delta}^{0} \frac{f(x+z)-f(x-0)}{z}\,dx \quad \text{und} \quad \int_{0}^{\delta} \frac{f(x+z)-f(x+0)}{z}\,dz \tag{12}$$

fordert. Dabei sind $f(x-0)$ und $f(x+0)$ der linksseitige bzw. rechtsseitige Grenzwert von f im Punkt x (es wird vorausgesetzt, daß x eine Unstetigkeitsstelle erster Art von f ist). In diesem Fall kann man nämlich die Differenz

$$S_n(x) - \frac{f(x+0)+f(x-0)}{z}$$

in der Form

$$\frac{1}{\pi}\int_{-\pi}^{0} [f(x+z)-f(x-0)]\,\frac{\sin\frac{2n+1}{2}z}{2\sin\frac{z}{2}}\,dz$$

$$+\frac{1}{\pi}\int_{0}^{\pi} [f(x+z)-f(x+0)]\,\frac{\sin\frac{2n+1}{2}z}{2\sin\frac{z}{2}}\,dz$$

darstellen, und diese Ausdrücke streben für $n \to \infty$ gegen Null, wenn die Integrale (12) existieren.

Hieraus folgt eine hinreichende Bedingung für die „globale“ Konvergenz der Fourierreihe, die gewöhnlich in den Analysisvorlesungen angegeben wird:

Es sei f eine beschränkte Funktion mit der Periode 2π. Ferner besitze f nur Unstetigkeitsstellen erster Art und in jedem Punkt eine rechtsseitige und eine linksseitige Ableitung[1]*). Dann konvergiert die Fourierreihe von f überall, und zwar an den Stetigkeitsstellen gegen $f(x)$ und an den Sprungstellen gegen $\big(f(x+0)+f(x-0)\big)/2$.*

2. Der Dirichletsche Kern $D_n(z)$, der in unseren Betrachtungen die entscheidende Rolle spielte, nimmt im Punkt $z = 0$ den Wert $(2n+1)/2\pi$ an und oszilliert für große n stark (Abb. 22). Deswegen liefert für große n nur eine beliebig kleine Umgebung des Punktes x den Hauptbeitrag zum Integral

$$\int_{-\pi}^{\pi} f(x+z)\, D_n(z)\, dz .$$

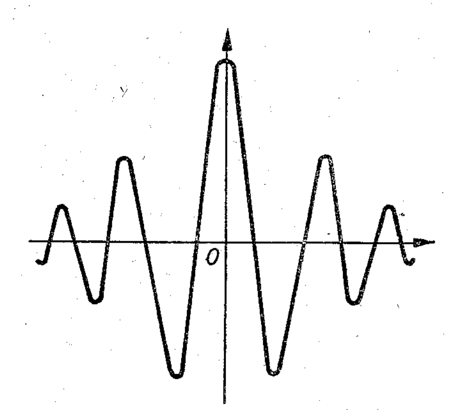


Abb. 22

Für Funktionen, die der Dinischen Bedingung genügen, strebt dieser Beitrag für $n \to \infty$ gegen $f(x)$. Faßt man die Dirichletschen Kerne $D_n(z)$ als Funktionale auf der Menge aller Funktionen auf, die in eine konvergente Fourierreihe entwickelbar sind, so konvergiert die Folge der $D_n(z)$ in einem gewissen Sinne gegen die δ-Funktion (δ-Distribution).

Es ist klar, daß die Folge $\{D_n\}$ im Sinne der gewöhnlichen Konvergenz keinen Grenzwert besitzt. Daher kann man auch nicht die Standardsätze über den Grenzübergang unter dem Integral zur Untersuchung des Integrals (7) anwenden.

---

[1]) In einer Unstetigkeitsstelle erster Art versteht man unter linksseitiger und rechtsseitiger Ableitung die Ausdrücke

$$\lim_{h\to 0+} \frac{f(x-h)-f(x-0)}{-h} \quad \text{bzw.} \quad \lim_{h\to 0+} \frac{f(x+h)-f(x+0)}{h}.$$

3. Die Dinische Bedingung, die die Konvergenz der Fourierreihe sichert, kann man auch durch andere Bedingungen ersetzen. Jedoch kann man die Dinische Bedingung in Satz 1 nicht einfach weglassen. Denn es gibt sogar stetige Funktionen, deren Fourierreihe in einigen Punkten divergiert, und es gibt integrierbare Funktionen, deren Fourierreihe überall divergiert (A. N. KOLMOGOROV). Schon 1915 stellte N. N. LUZIN folgendes Problem: Gibt es im Raum L_2 Funktionen, deren Fourierreihe auf einer Menge mit positivem Maß divergiert? Wie L. CARLESON gezeigt hat (1966), existieren solche Funktionen nicht.

Die Existenz stetiger Funktionen, deren Fourierreihe nicht in allen Punkten konvergiert, ergibt sich leicht aus dem allgemeinen Satz über die schwache Konvergenz von Funktionalen. Zum Nachweis vermerken wir zunächst, daß

$$\int_{-\pi}^{\pi} |D_n(z)|\, dz \to \infty \quad \text{für } n \to \infty \tag{13}$$

ist. Um das zu zeigen, betrachten wir den Zähler des Bruches

$$|D_n(z)| = \frac{\left|\sin \frac{2n+1}{2} z\right|}{2\pi \left|\sin \frac{z}{2}\right|}.$$

Er nimmt in den Punkten z_k mit

$$\frac{2n+1}{2} z_k = \left(k + \frac{1}{2}\right)\pi \quad (k = 0, 1, \ldots, n) \tag{14}$$

den Wert 1 an. Auf jedem Intervall

$$|z - z_k| < \frac{2\pi}{3(2n+1)} \tag{15}$$

ist dann $\left|\sin \frac{2n+1}{2} z\right|$ nicht kleiner als $\frac{1}{2}$. Für den Wert von $\sin \frac{z}{2}$ auf dem k-ten Intervall $(k = 0, 1, \ldots, n)$ gilt dabei die Abschätzung

$$\sin \frac{z}{2} < \frac{z}{2} < \frac{1}{2}\left(\frac{2k+1}{2}\pi + \frac{\pi}{3}\right)\left(\frac{2n+1}{2}\right)^{-1} < \frac{k+1}{2n+1}\pi.$$

Daher ist das Integral von $|D_n(z)|$ bereits auf den durch (15) definierten Abschnitten größer als die Summe

$$\frac{1}{2\pi}\sum_{k=0}^{n} \frac{1}{2} \cdot \frac{1}{\frac{k+1}{2n+1}\pi} \cdot \frac{4\pi}{3(2n+1)} = \frac{1}{3\pi}\sum_{k=0}^{n} \frac{1}{k+1},$$

die für $n \to \infty$ gegen ∞ strebt. Hieraus ergibt sich die Beziehung (13). Das bedeutet, daß die Normen der Funktionale D_n auf dem Raum der stetigen Funktionen nicht gleichmäßig beschränkt sind. Nach dem Satz über die schwache Konvergenz von Funktionalen kann die Folge $\{D_n\}$ dann aber auf diesem Raum auch nicht schwach konvergent sein. Es gibt also stetige Funktionen f, für die der Grenzwert

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} D_n(x) f(x) \, dx$$

nicht existiert.

8.1.2. Bedingungen für die gleichmäßige Konvergenz der Fourierreihe. Wir haben hinreichende Bedingungen für die Konvergenz der Fourierreihe einer Funktion f in jedem Punkt aufgestellt. Die Klasse der Funktionen, die diesen Bedingungen genügen, ist sehr groß. Sogar die Stetigkeit ist keinesfalls notwendig für die Darstellbarkeit einer Funktion durch ihre überall konvergente Fourierreihe. Die Situation ändert sich, wenn wir uns dafür interessieren, wann die Fourierreihe gleichmäßig konvergiert. Besitzt nämlich die Funktion f auch nur eine Unstetigkeitsstelle, so kann die Fourierreihe von f nicht gleichmäßig gegen f konvergieren, da die Summe einer gleichmäßig konvergenten Reihe stetiger Funktionen wieder stetig ist. Somit ist die Stetigkeit einer Funktion eine notwendige (aber natürlich nicht hinreichende) Bedingung für die gleichmäßige Konvergenz ihrer Fourierreihe. Der folgende Satz liefert eine einfache hinreichende Bedingung.

Satz 2. *Ist die Funktion f mit der Periode 2π absolut stetig und gehört ihre Ableitung f' zu $L_2[-\pi, \pi]$, so konvergiert die Fourierreihe von f auf der ganzen reellen Achse gleichmäßig gegen f.*

Beweis. Wir bezeichnen die Fourierkoeffizienten der Funktion f' mit a_n' und b_n'. Da die Funktion f absolut stetig ist, darf man das Integral

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx$$

partiell integrieren. Danach ergibt sich

$$\begin{aligned} a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx \\ &= \frac{1}{\pi} f(x) \frac{\sin nx}{n} \Big|_{-\pi}^{\pi} - \frac{1}{n\pi} \int_{-\pi}^{\pi} f'(x) \sin nx \, dx \\ &= -\frac{b_n'}{n}. \end{aligned}$$

Analog erhält man

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx \, dx = \frac{a_n'}{n}.$$

Folglich ist

$$\frac{|a_0|}{2} + \sum_{n=1}^{\infty} |a_n| + |b_n| = \frac{|a_0|}{2} + \sum_{n=1}^{\infty} \frac{|b_n'|}{n} + \frac{|a_n'|}{n}. \tag{16}$$

Diese Reihe konvergiert wegen

$$\frac{|b_n'|}{n} \leqq \frac{1}{2} \left(b_n'^2 + \frac{1}{n^2} \right), \quad \frac{|a_n'|}{n} \leqq \frac{1}{2} \left(a_n'^2 + \frac{1}{n^2} \right)$$

und der Besselschen Ungleichung

$$\sum_{n=1}^{\infty} (b_n'^2 + a_n'^2) \leqq \|f'\|^2 < \infty.$$

Die Reihe (16) ist aber offenbar eine Majorante für die Fourierreihe von f. Nach dem Weierstraßschen Kriterium konvergiert daher die Fourierreihe von f (absolut und) gleichmäßig gegen f.

Es bleibt zu zeigen, daß die Summe dieser Reihe tatsächlich f ist. Dazu bezeichnen wir die Summe der Fourierreihe von f mit φ. Dann besitzt φ dieselben Fourierkoeffizienten wie f. Hieraus ergibt sich $f = \varphi$ wegen der Stetigkeit der beiden Funktionen.

Man kann für die gleichmäßige Konvergenz der Fourierreihe auch eine andere Bedingung angeben, die der Dinischen Bedingung analog ist.

Satz 3. *Die integrierbare Funktion f sei auf einer Menge $E \subset [-\pi, \pi]$ beschränkt, und die Dinische Bedingung sei gleichmäßig auf E erfüllt: Für jedes $\varepsilon > 0$ existiere also ein $\delta > 0$ mit*

$$\int_{-\delta}^{\delta} \frac{|f(x+z) - f(x)|}{|z|} \, dz < \varepsilon,$$

und zwar gleichzeitig für alle $x \in E$. Dann konvergiert die Fourierreihe von f auf E gleichmäßig gegen f.

Der Beweis dieses Satzes beruht auf einem Lemma, das eine Verschärfung von Lemma 1 (vgl. 8.1.1.) darstellt.

Lemma 2. *B sei eine präkompakte Menge im Raum $L_1[-\pi, \pi]$. Dann gibt es zu jedem $\varepsilon > 0$ eine Zahl $N = N(\varepsilon)$ mit*

$$\left| \int_a^b f(t) \sin \lambda t \, dt \right| < \varepsilon$$

für $\lambda \geqq N(\varepsilon)$, gleichzeitig für alle $f \in B$.

Zum Beweis des Lemmas wählen wir in B ein endliches $\varepsilon/2$-Netz $\varphi_1, \ldots, \varphi_k$ und bestimmen ein N mit

$$\left| \int_a^b \varphi_i(t) \sin \lambda t \, dt \right| < \frac{\varepsilon}{2} \quad \text{für} \quad \lambda \geqq N, \; i = 1, 2, \ldots, k.$$

Ist nun f eine beliebige Funktion aus B, so gilt

$$\|f - \varphi_i\| < \frac{\varepsilon}{2}$$

für ein gewisses i und folglich

$$\left| \int_a^b f(t) \sin \lambda t \, dt \right| \leqq \left| \int_a^b \varphi_i(t) \sin \lambda t \, dt \right| + \left| \int_a^b (f - \varphi_i) \sin \lambda t \, dt \right| < \varepsilon.$$

Damit ist das Lemma bewiesen.

Die Verwendung dieses Lemmas zum Beweis von Satz 3 beruht auf der leicht beweisbaren Tatsache, daß die Menge der Funktionen

$$\varphi_x(t) = \frac{f(x + t) - f(x)}{t}$$

präkompakt ist. Den ausführlichen Beweis überlassen wir dem Leser.

Bis jetzt haben wir Funktionen auf dem Intervall $[-\pi, \pi]$ betrachtet. Es ist aber klar, daß alles bisher Gesagte auch für Funktionen auf einem beliebigen Intervall der Länge $2l$ gilt.

Auch für den Fall mehrerer Veränderlicher kann man hinreichende Bedingungen sowohl für die Konvergenz der Fourierreihe in jedem Punkt als auch für die gleichmäßige Konvergenz der Fourierreihe formulieren. Wir werden uns damit jedoch nicht aufhalten.

8.2. Der Fejérsche Satz

8.2.1. Der Fejérsche Satz. Es sei f eine stetige Funktion auf der reellen Achse mit der Periode 2π. Diese Funktion wird durch ihre Fourierreihe

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \tag{1}$$

eindeutig bestimmt. Denn sind f_1 und f_2 zwei stetige Funktionen mit denselben Fourierkoeffizienten, so ist die stetige Funktion $f_1 - f_2$ fast überall Null und folglich identisch Null. Da jedoch die Fourierreihe einer stetigen Funktion nicht notwendig konvergiert, können wir die Funktion f im allgemeinen nicht unmittelbar durch Summation ihrer Fourierreihe erhalten. Eine Möglichkeit, stetige Funktionen durch ihre Fourierreihe darzustellen, liefert ein Satz, der 1905 von Fejér bewiesen wurde.

Es seien

$$S_k(x) = \frac{a_0}{2} + \sum_{j=1}^{k} (a_j \cos jx + b_j \sin jx), \qquad k = 1, 2, \ldots, \tag{2}$$

die Partialsummen der Fourierreihe der Funktion f. Damit setzen wir

$$\sigma_n(x) = \frac{S_0(x) + S_1(x) + \cdots + S_{n-1}(x)}{n}. \tag{3}$$

Diese arithmetischen Mittel σ_n der Partialsummen S_k heißen *Fejérsche Summen* der Funktion f.

Satz 1 (Fejér). *Ist f eine stetige Funktion mit der Periode 2π, so konvergiert die Folge $\{\sigma_n\}$ der Fejérschen Summen von f auf der ganzen reellen Achse gleichmäßig gegen f.*

Beweis. Wir benutzen die in 8.1. erhaltene Integraldarstellung

$$S_k(x) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x+z) \frac{\sin \frac{2k+1}{2} z}{2 \sin \frac{z}{2}} \, dz$$

für die Partialsummen der Fourierreihe von f. Setzen wir diesen Ausdruck in die Gleichung (3) ein, so erhalten wir

$$\sigma_n(x) = \frac{1}{2n\pi} \int_{-\pi}^{\pi} \left\{ \sum_{k=0}^{n-1} \frac{\sin \frac{2k+1}{2} z}{\sin \frac{z}{2}} \right\} \cdot f(x+z) \, dz.$$

Daraus folgt mit Hilfe der Formel[1])

$$\sum_{k=0}^{n-1} \sin (2k+1) z = \frac{\sin^2 nz}{\sin z}$$

---

[1]) Diese Formel erhält man leicht durch Summation der Gleichungen

$$2 \cdot \sin (2k+1) z \cdot \sin z = \cos 2kz - \cos 2(k+1) z \qquad (k = 0, 1, \ldots, n-1).$$

die Darstellung von $\sigma_n(x)$ in Form eines sogenannten *Fejérschen Integrals*,

$$\sigma_n(x) = \frac{1}{2n\pi} \int_{-\pi}^{\pi} \left(\frac{\sin n \frac{z}{2}}{\sin \frac{z}{2}} \right)^2 f(x+z)\, dz. \tag{4}$$

Dabei wird der Ausdruck

$$\Phi_n(z) = \frac{1}{2n\pi} \left(\frac{\sin n \frac{z}{2}}{\sin \frac{z}{2}} \right)^2 \tag{5}$$

Fejérscher Kern genannt. Formel (4) kann man damit in der Gestalt

$$\sigma_n(x) = \int_{-\pi}^{\pi} f(x+z)\, \Phi_n(z)\, dz \tag{6}$$

schreiben. Wir müssen also zeigen, daß dieser Ausdruck für $n \to \infty$ gleichmäßig gegen $f(x)$ strebt.

Dazu vermerken wir vorbereitend die folgenden Eigenschaften des Fejérschen Kerns:

1. $\Phi_n(z) \geqq 0$,

2. $\int_{-\pi}^{\pi} \Phi_n(z)\, dz = 1$,

3. für jedes feste $\delta > 0$ und $n \to \infty$ gilt

$$\int_{-\pi}^{\delta} \Phi_n(z)\, dz = \int_{\delta}^{\pi} \Phi_n(z)\, dz = \eta_n(\delta) \to 0.$$

Die erste dieser Eigenschaften ist trivial. Die zweite folgt aus Gleichung (6), denn für $f(x) \equiv 1$ ist $\sigma_n(x) \equiv 1$ für alle n. Die dritte Eigenschaft ergibt sich unmittelbar daraus, daß für $\delta < z \leqq \pi$ die Abschätzung $\sin \frac{z}{2} \geqq \frac{2\delta}{\pi}$ und folglich auch die Ungleichung

$$\left(\frac{\sin n \frac{z}{2}}{\sin \frac{z}{2}} \right)^2 \leqq \left(\frac{\pi}{2\delta} \right)^2$$

gültig ist. Wenn man diese Eigenschaften des Fejérschen Kerns berücksichtigt, läßt sich der Satz jetzt unschwer beweisen.

Die Funktion f ist periodisch und stetig; somit ist sie auf der ganzen reellen Achse beschränkt und gleichmäßig stetig. Es existiert also eine Konstante M mit

$$|f(x)| \leqq M \tag{7}$$

für alle x, und zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so daß

$$|f(x'') - f(x')| < \frac{\varepsilon}{2} \quad \text{für } |x'' - x'| < \delta \tag{8}$$

ist. Zum Beweis des Satzes müssen wir die Differenz

$$f(x) - \sigma_n(x) = \int_{-\pi}^{\pi} [f(x) - f(x+z)]\, \Phi_n(z)\, dz$$

abschätzen, die man als Summe der drei Integrale

$$I_- = \int_{-\pi}^{-\delta} \{f(x) - f(x+z)\}\, \Phi_n(z)\, dz,$$

$$I_0 = \int_{-\delta}^{\delta} \{f(x) - f(x+z)\}\, \Phi_n(z)\, dz,$$

$$I_+ = \int_{\delta}^{\pi} \{f(x) - f(x+z)\}\, \Phi_n(z)\, dz$$

darstellen kann. Aus (7) und (8) ergeben sich unmittelbar die Abschätzungen

$$|I_-| \leqq 2M\eta_n(\delta),$$

$$|I_+| \leqq 2M\eta_n(\delta),$$

$$|I_0| \leqq \frac{\varepsilon}{2} \int_{-\delta}^{\delta} \Phi_n(z)\, dz < \frac{\varepsilon}{2}.$$

Wählen wir jetzt n_0 so groß, daß bei gegebenem δ für $n \geqq n_0$ die Ungleichung

$$2M\eta_n(\delta) < \frac{\varepsilon}{4}$$

erfüllt ist, so gilt

$$|f(x) - \sigma_n(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \varepsilon.$$

Daraus folgt die Behauptung des Satzes, denn ε war beliebig.

Wir weisen darauf hin, daß wir beim Beweis nur die Eigenschaften 1 bis 3 des Fejérschen Kerns ausgenutzt haben. Das gestattet, verschiedene Verallgemeinerungen von Satz 1 zu erhalten (siehe insbesondere 8.3.).

8.2.2. Die Vollständigkeit des trigonometrischen Systems. Der Weierstraßsche Approximationssatz. Aus dem Fejérschen Satz ergibt sich die Vollständigkeit des trigonometrischen Systems im Raum $L_2[-\pi, \pi]$. Nach diesem Satz ist nämlich jede stetige Funktion der Grenzwert der gleichmäßig (und folglich auch im Mittel) konvergenten Folge $\{\sigma_n(x)\}$ von trigonometrischen Polynomen. Es bleibt nur noch zu bemerken, daß die stetigen Funktionen in L_2 dicht sind.

Man kann den Fejérschen Satz als eine Verschärfung des Weierstraßschen Approximationssatzes ansehen, nach dem jede stetige periodische Funktion durch trigonometrische Polynome approximiert werden kann. Während der Weierstraßsche Approximationssatz nur die *Existenz* einer gleichmäßig konvergenten, approximierenden Folge von trigonometrischen Polynomen sichert, gibt der Fejérsche Satz eine solche Folge explizit an, nämlich die Folge der Fejérschen Summen (3).

Aus dem Satz von WEIERSTRASS über die gleichmäßige Approximierbarkeit einer stetigen periodischen Funktion durch trigonometrische Polynome ergibt sich leicht ein zweiter Satz von WEIERSTRASS über die Approximierbarkeit einer beliebigen stetigen Funktion auf einem Intervall $[a, b]$ durch algebraische Polynome. Dazu betrachten wir eine stetige Funktion $f(x)$ auf dem Intervall $[a, b]$. Durch die Substitution $x = \frac{t(b-a)}{\pi} + a$ erhalten wir daraus eine Funktion $\varphi(t)$ auf dem Intervall $[0, \pi]$. Diese Funktion setzen wir zunächst durch die Gleichung $\varphi(-t) = \varphi(t)$ auf das Intervall $[-\pi, 0)$ und anschließend periodisch auf die gesamte reelle Achse fort. Dann gibt es ein trigonometrisches Polynom T_n mit

$$|T_n(t) - \varphi(t)| < \frac{\varepsilon}{2} \quad \text{für alle } t.$$

Ferner kann man jedes trigonometrische Polynom in seine Taylorreihe entwickeln, die auf jedem endlichen Intervall gleichmäßig konvergiert. P_m sei eine Partialsumme der Taylorreihe von T_n mit

$$|T_n(t) - P_m(t)| < \frac{\varepsilon}{2} \quad \text{für } 0 \leqq t \leqq \pi.$$

Dann ist also

$$|\varphi(t) - P_m(t)| < \varepsilon \quad \text{für } 0 \leqq t \leqq \pi.$$

Setzen wir nun wieder $t = \frac{x-a}{b-a}\pi$, so ist $Q_m(x) = P_m\left(\frac{x-a}{b-a}\pi\right) = P_m(t)$ ein Polynom mit

$$|f(x) - Q_m(x)| < \varepsilon \quad \text{für } a \leqq x \leqq b.$$

8.2.3. Der Fejérsche Satz für den Raum L_1. Im Fejérschen Satz besteht eine gewisse Symmetrie zwischen der Voraussetzung und der Behauptung des Satzes. Aus der Zugehörigkeit der Funktion f zum Raum $C[-\pi, \pi]$ folgt, daß die zugehörige Fejérsche Summe in der Metrik desselben Raumes $C[-\pi, \pi]$ gegen f konvergiert. Analoge Sätze kann man auch für andere Funktionen-

räume erhalten, insbesondere für den Raum $L_1[-\pi, \pi]$. Es gilt also der folgende Satz, den man dann als Fejérschen Satz für integrierbare Funktionen bezeichnet:

Ist f eine auf dem Intervall $[-\pi, \pi]$ integrierbare Funktion, dann konvergieren die Fejérschen Summen von f in der Norm des Raumes $L_1[-\pi, \pi]$ gegen f.

Den Beweis dieser Behauptung kann man mit Hilfe von ähnlichen Betrachtungen wie in 8.2.1. erhalten. Wir wollen ihn hier nicht durchführen. Jedoch verweisen wir auf folgende wichtige Tatsache, die sich aus dem Fejérschen Satz für integrierbare Funktionen ergibt. *Jede integrierbare Funktion ist durch ihre Fourierkoeffizienten eindeutig* (bis auf Äquivalenz) *bestimmt.*

Um das zu zeigen, betrachten wir zwei integrierbare Funktionen f und g mit denselben Fourierkoeffizienten. Dann verschwinden alle Fourierkoeffizienten der Funktion $f-g$. Folglich sind auch alle Fejérschen Summen von $f-g$ identisch 0. Ihr Grenzwert in L_1, d. h. die Funktion $f-g$, muß dann aber fast überall 0 sein.

8.3. Das Fouriersche Integral

8.3.1. Hauptsatz. In 8.1. haben wir Bedingungen aufgestellt, unter denen man eine periodische Funktion in eine Fourierreihe entwickeln, d. h. als Überlagerung harmonischer Schwingungen darstellen kann. Wir wollen jetzt versuchen, dieses Resultat auf nichtperiodische Funktionen zu übertragen. Wie wir sehen werden, ist eine solche Darstellung bereits unter ziemlich allgemeinen Bedingungen möglich, allerdings nicht mehr mit Hilfe einer Reihe, sondern mit Hilfe eines Integrals. Dieses Integral nennt man *Fouriersches Integral.*

Es sei f eine Funktion, die auf jedem endlichen Intervall in eine Fourierreihe entwickelbar ist. Dazu nehmen wir zum Beispiel an, daß f auf jedem endlichen Intervall und in jedem Punkt der Dinischen Bedingung genügt. Wir betrachten f etwa auf dem Intervall $[-l, l]$ und entwickeln diese Funktion in bezug auf dieses Intervall in ihre Fourierreihe:

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos \frac{k\pi}{l} x + b_k \sin \frac{k\pi}{l} x. \tag{1}$$

Hier ersetzen wir überall a_0, a_k und b_k durch die Ausdrücke

$$a_0 = \frac{1}{l} \int_{-l}^{l} f(t)\, dt,$$

$$a_k = \frac{1}{l} \int_{-l}^{l} f(t) \cos \frac{k\pi}{l} t\, dt,$$

$$b_k = \frac{1}{l} \int_{-l}^{l} f(t) \sin \frac{k\pi}{l} t\, dt$$

Dann erhalten wir

$$f(x) = \frac{1}{2l} \int_{-l}^{l} f(t)\, dt$$

$$+ \sum_{k=1}^{\infty} \left[\frac{1}{l} \int_{-l}^{l} f(t) \cos \frac{k\pi}{l} x \cos \frac{k\pi}{l} t\, dt + \frac{1}{l} \int_{-l}^{l} f(t) \sin \frac{k\pi}{l} x \sin \frac{k\pi}{l} t \right] dt$$

$$= \frac{1}{2l} \int_{-l}^{l} f(t)\, dt$$

$$+ \frac{1}{l} \sum_{k=1}^{\infty} \int_{-l}^{l} f(t) \left[\cos \frac{k\pi}{l} x \cos \frac{k\pi}{l} t + \sin \frac{k\pi}{l} x \sin \frac{k\pi}{l} t \right] dt,$$

d. h.

$$f(x) = \frac{1}{2l} \int_{-l}^{l} f(t)\, dt + \frac{1}{\pi} \sum_{k=1}^{\infty} \frac{\pi}{l} \int_{-l}^{l} f(t) \cos \frac{k\pi}{l} (t - x)\, dt. \tag{2}$$

Wir wollen von der Funktion f jetzt noch zusätzlich voraussetzen, daß sie auf der ganzen reellen Achse absolut integrierbar ist, d. h.

$$\int_{-\infty}^{\infty} |f(t)|\, dt < \infty. \tag{3}$$

Gehen wir nun in der Gleichung (2) (zunächst rein formal) zum Grenzwert für $l \to \infty$ über, so strebt der erste Summand auf der rechten Seite wegen (3) gegen Null. Den zweiten Summanden kann man als (auf das Intervall $[0, \infty)$ bezogene) Partialsumme des Integrals

$$\int_{0}^{\infty} F(\lambda)\, d\lambda$$

der Funktion

$$F(\lambda) = \int_{-l}^{l} f(t) \cos \lambda(t - x)\, dt$$

auffassen, wenn man $\lambda_k = k\pi/l$ und $\Delta\lambda = \pi/l$ setzt. Damit führt der formale Grenzübergang $l \to \infty$ in (2) zu der Gleichung

$$f(x) = \frac{1}{\pi} \int_{0}^{\infty} d\lambda \int_{-\infty}^{\infty} f(t) \cos \lambda(t - x)\, dt. \tag{4}$$

Das ist gerade die gesuchte Darstellung. Mit den Bezeichnungen

$$a_\lambda = \frac{1}{\pi} \int_{-\infty}^{\infty} f(t) \cos \lambda t \, dt,$$

$$b_\lambda = \frac{1}{\pi} \int_{-\infty}^{\infty} f(t) \sin \lambda t \, dt$$

kann man Gleichung (4) in Analogie zur Fourierreihe auch in der Form

$$f(x) = \int_0^\infty [a_\lambda \cos \lambda x + b_\lambda \sin \lambda x) \, d\lambda \tag{5}$$

schreiben.

Wir haben Gleichung (4), die sogenannte *Fouriersche Formel*, mit Hilfe eines formalen Grenzübergangs erhalten. Man könnte die Richtigkeit dieses Grenzübergangs (unter den obigen Voraussetzungen über die Funktion f) nachweisen. Es ist jedoch einfacher, unmittelbar Gleichung (4) zu zeigen.

Satz 1. *Ist die Funktion f auf der ganzen reellen Achse absolut integrierbar und genügt sie im Punkt x der Dinischen Bedingung, so gilt*

$$f(x) = \frac{1}{\pi} \int_0^\infty d\lambda \int_{-\infty}^{\infty} f(t) \cos \lambda (t - x) \, dt.$$

Beweis. Wir führen die Bezeichnung

$$I(A) = \frac{1}{\pi} \int_0^A d\lambda \int_{-\infty}^{\infty} f(t) \cos \lambda (t - x) \, dt \tag{6}$$

ein. Dann müssen wir zeigen, daß $\lim_{A \to \infty} I(A)$ existiert und gleich $f(x)$ ist. Ersetzt man den Integranden in (6) durch seinen Absolutbetrag, so existiert zunächst das innere Integral wegen der absoluten Integrierbarkeit von f für alle λ. Da diese Integrale gleichmäßig beschränkt sind, existiert dann auch das iterierte Integral. Nach dem Satz von Fubini können wir daher in (6) die Reihenfolge der Integration vertauschen,

$$I(A) = \frac{1}{\pi} \int_{-\infty}^{\infty} dt \int_0^A f(t) \cos \lambda (t - x) \, d\lambda = \frac{1}{\pi} \int_{-\infty}^{\infty} f(t) \frac{\sin A(t - x)}{t - x} \, dt.$$

Durch die Substitution $t - x = z$ geht dieses Integral in

$$I(A) = \frac{1}{\pi} \int_{-\infty}^{\infty} f(x + z) \frac{\sin Az}{z} \, dz \tag{7}$$

über. Die wohlbekannte Beziehung

$$\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin Az}{z}\,dz = 1 \qquad (A > 0)$$

gestattet dann für die Differenz $I(A) - f(x)$ die Darstellung

$$I(A) - f(x) = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{f(x+z) - f(x)}{z}\sin Az\,dz. \tag{8}$$

Das rechts stehende Integral zerlegen wir in drei Summanden:

$$\begin{aligned} I(A) - f(x) = {} & \frac{1}{\pi}\int_{-N}^{N}\frac{f(x+z) - f(x)}{z}\sin Az\,dz \\ & + \frac{1}{\pi}\int_{|z|\geqq N}\frac{f(x+z)}{z}\sin Az\,dz \\ & - \frac{f(x)}{\pi}\int_{|z|\geqq N}\frac{\sin Az}{z}\,dz. \end{aligned}$$

Der zweite und der dritte Summand auf der rechten Seite werden für hinreichend großes N kleiner als $\varepsilon/3$, da die entsprechenden Integrale konvergent sind. Der erste Summand rechts strebt (bei festem N) für $A \to \infty$ gegen Null (nach 8.1., Lemma 1, und der Dinischen Bedingung). Somit erhalten wir

$$\lim_{A\to\infty}\bigl(I(A) - f(x)\bigr) = 0,$$

was zu zeigen war.

8.3.2. Das Fouriersche Integral in komplexer Form. Das innere Integral in der Integralformel (4) ist eine gerade Funktion von λ. Daher kann man diese Formel auch in der Gestalt

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}d\lambda\int_{-\infty}^{\infty}f(t)\cos\lambda(t - x)\,dt \tag{9}$$

schreiben. Ferner folgt aus der absoluten Integrierbarkeit von f, daß das Integral $\int_{-\infty}^{\infty} f(t) \sin \lambda(t - x)\, dt$ existiert. Es stellt eine ungerade Funktion von λ dar, also ist

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} d\lambda \int_{-\infty}^{\infty} f(t) \sin \lambda(t - x)\, dt = 0 \tag{10}$$

(wenn das Integral über λ im Sinne der Hauptwerts verstanden wird, d. h. als $\lim_{N\to\infty} \int_{-N}^{N}$).
Wenn wir (10) mit $-i$ multiplizieren und zu (9) addieren, so erhalten wir

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\lambda \int_{-\infty}^{\infty} f(t)\, e^{-i\lambda(t-x)}\, dt.$$

Das ist die *Fouriersche Integralformel in komplexer Form.*

8.4. Die Fouriertransformation. Eigenschaften und Anwendungen

8.4.1. Die Fouriertransformierte und die Umkehrformel. Die Fouriersche Integralformel kann man in zwei Gleichungen zerlegen. Setzen wir

$$g(\lambda) = \int_{-\infty}^{\infty} f(t)\, e^{-i\lambda t}\, dt, \tag{1}$$

so ist

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} g(\lambda)\, e^{i\lambda x}\, d\lambda. \tag{2}$$

Dabei hat Formel (1) für beliebige integrierbare Funktionen f einen Sinn. Somit wird durch (1) jedem $f \in L_1(-\infty, \infty)$ eine bestimmte Funktion g auf der Zahlengeraden zugeordnet. Diese Funktion g heißt *Fouriertransformierte* der Ausgangsfunktion f. Formel (2), die die Funktion f durch ihre Fouriertransformierte ausdrückt, heißt *Fouriersche Umkehrformel.* Hierbei fällt die Ähnlichkeit zwischen den Formeln (1) und (2) auf. Die zweite Formel unterscheidet sich von der ersten nur durch das Vorzeichen im Exponenten und den Faktor $1/2\pi$ vor dem Integral. Man könnte hier noch größere Symmetrie erreichen, indem man g durch

$$g(\lambda) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)\, e^{-i\lambda x}\, dx \tag{1'}$$

definiert. Dann würde die Umkehrformel die Gestalt

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(\lambda)\, e^{i\lambda x}\, d\lambda \tag{2'}$$

annehmen, d. h., es bliebe nur der Unterschied im Vorzeichen des Exponenten.

Bei aller äußerlichen Ähnlichkeit sind jedoch die Formeln (1) und (2) ihrem Wesen nach verschieden: Das erste der Integrale existiert (wegen $f \in L_1(-\infty, \infty)$) im gewöhnlichen Sinne, das zweite im allgemeinen nur im Sinne des Hauptwerts. Außerdem ist (1) die Definitionsgleichung einer Funktion g, während (2) als Umschrift der Fourierschen Integralformel die Behauptung enthält, daß das rechts stehende Integral gleich der Ausgangsfunktion f ist. Wie wir oben gesehen haben, muß man zum Nachweis dieser Gleichung außer der Integrierbarkeit noch zusätzliche Bedingungen über f voraussetzen, etwa die Dinische Bedingung.

Bemerkung. Wir haben für jede Funktion $f \in L_1(-\infty, \infty)$ die Fouriertransformierte g definiert. Dabei ließ sich die Funktion f mit Hilfe der Umkehrformel durch die Fouriertransformierte g ausdrücken, wenn f in jedem Punkt der Dinischen Bedingung genügt. Es liegt also ein vollkommen analoger Sachverhalt vor wie bei den Fourierreihen. Dort wurden die Fourierkoeffizienten

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\, e^{-inx}\, dx$$

für jedes $f \in L_1[-\pi, \pi]$ definiert, die Konvergenz der Fourierreihe

$$\sum_{n=-\infty}^{\infty} c_n e^{inx}$$

(die hier die Rolle der Umkehrformel spielt) konnte jedoch nur unter bestimmten zusätzlichen Bedingungen (Dinische Bedingung) garantiert werden. Gleichzeitig gilt für die Fouriertransformierte (wie auch für die Fourierreihe; vgl. den Schluß von 8.2.) Folgendes: *Wenn für die Funktion* $f \in L_1(-\infty, \infty)$

$$\int_{-\infty}^{\infty} f(x)\, e^{-i\lambda x}\, dx \equiv 0$$

ist, so gilt $f(x) = 0$ *fast überall.*

Aus der obigen Voraussetzung ergibt sich die Beziehung

$$\int_{-\infty}^{\infty} f(x + t)\, e^{-i\lambda x}\, dx = 0$$

für alle reellen Zahlen t und λ. Wir setzen jetzt

$$\varphi(x) = \int_0^{\xi} f(x + t)\, dt,$$

wobei ξ eine beliebige, aber feste reelle Zahl ist. Diese Funktion φ gehört wie f zu $L_1(-\infty, \infty)$ und genügt der Bedingung

$$\int_{-\infty}^{\infty} \varphi(x)\, e^{-i\lambda x}\, dx = 0$$

für alle reellen Zahlen λ. Das erhält man leicht, wenn man den Satz von FUBINI anwendet und die Voraussetzung über die Funktion f ausnutzt. Unschwer sieht man nun, daß die Funktion φ auf jedem endlichen Intervall absolut stetig ist. Folglich besitzt sie fast überall eine endliche Ableitung. Insbesondere genügt diese Funktion also fast überall der Dinischen Bedingung. Auf Grund von 8.3., Satz 1, verschwindet sie daher fast überall, denn ihre Fouriertransformierte ist identisch 0. Wegen der Stetigkeit von φ ist dann sogar $\varphi(x) \equiv 0$. Daraus ergibt sich insbesondere, daß für alle reellen Zahlen ξ

$$\int_0^{\xi} f(t)\, dt = 0$$

ist. Folglich gilt fast überall $f(x) = 0$.

Wir wollen jetzt einige Beispiele betrachten.

1. Es sei $f(x) = e^{-\gamma|x|}$, $\gamma > 0$. Wir wollen die Fouriertransformierte dieser Funktion bestimmen. Es ist

$$\begin{aligned} g(\lambda) &= \int_{-\infty}^{\infty} e^{-\gamma|x|}\, e^{-i\lambda x}\, dx \\ &= \int_{-\infty}^{\infty} e^{-\gamma|x|}\, (\cos \lambda x - i \sin \lambda x)\, dx \\ &= 2 \int_0^{\infty} e^{-\gamma|x|} \cos \lambda x\, dx, \end{aligned}$$

und nach zweimaliger partieller Integration ergibt sich daraus

$$g(\lambda) = \frac{2\gamma}{\lambda^2 + \gamma^2}.$$

2. Es sei

$$f(x) = \begin{cases} 1 & \text{für } |x| \leqq a, \\ 0 & \text{für } |x| > a. \end{cases}$$

Dann gilt

$$g(\lambda) = \int_{-\infty}^{\infty} f(x)\, e^{-i\lambda x}\, dx = \int_{-a}^{a} e^{-i\lambda x}\, dx = \frac{e^{i\lambda a} - e^{-i\lambda a}}{i\lambda} = \frac{2 \sin \lambda a}{\lambda}.$$

(Man beachte, daß die Funktion g hier n i c h t zu $L_1(-\infty, \infty)$ gehört.)

3. Es sei $f(x) = \dfrac{1}{x^2 + a^2}$. Dann ist

$$g(\lambda) = \int_{-\infty}^{\infty} e^{-i\lambda x} \frac{dx}{x^2 + a^2}. \tag{3}$$

Dieses Integral berechnet man am einfachsten mit Hilfe der Residuentheorie. Zunächst sei $\lambda > 0$. Wir fügen zur reellen Achse, über die sich das Integral (3) erstreckt, einen Halbkreis mit unendlich großem Radius in der unteren Halbebene hinzu (d. h. dort, wo die Exponentialfunktion $e^{-i\lambda x}$ gegen Null strebt). Dann ist das Integral (3) gleich der Summe der Residuen des Integranden in der unteren Halbebene, multipliziert mit $2\pi i$. In der unteren Halbebene besitzt die Funktion $\dfrac{e^{-i\lambda x}}{x^2 + a^2}$ einen Pol erster Ordnung im Punkt $x = -ai$. Das Residuum in diesem Punkt findet man nach dem bekannten Residuensatz: *Ist* $f(z) = \dfrac{\varphi(z)}{\psi(z)}$ *und ist* $\varphi(a) \neq 0$, *während* $\psi(z)$ *im Punkt* $z = a$ *eine Nullstelle erster Ordnung besitzt, so ist das Residuum der Funktion* f *im Punkt* a *gleich* $\dfrac{\varphi(a)}{\psi'(a)}$. Danach erhalten wir in unserem Fall

$$g(\lambda) = \frac{1}{2\pi i} \frac{e^{-a\lambda}}{-2ai} = \frac{e^{-a\lambda}}{4\pi a} \quad \text{für } \lambda > 0.$$

Für $\lambda < 0$ erhält man analog

$$g(\lambda) = \frac{e^{a\lambda}}{4\pi a}$$

(wenn man anstelle der unteren die obere Halbebene betrachtet). Somit gilt schließlich

$$g(\lambda) = \frac{e^{-a|\lambda|}}{4\pi a} \quad (-\infty < \lambda < \infty).$$

Übrigens kann man dieses Resultat auch sofort mit Hilfe der Umkehrformel erhalten, wenn man Beispiel 1 und 8.3., Satz 1, ausnutzt.

4. Setzen wir $f(x) = e^{-ax^2}$, so ist

$$g(\lambda) = \int_{-\infty}^{\infty} e^{-ax^2} e^{-i\lambda x}\, dx. \tag{4}$$

Hier steht unter dem Integral eine analytische Funktion, die in der endlichen Ebene keine Singularitäten besitzt. Außerdem strebt diese Funktion längs jeder zur x-Achse parallelen Geraden gegen Null. Nach dem Cauchyschen Integralsatz ändert daher das

Integral (4) seinen Wert nicht, wenn es nicht längs der reellen Geraden, sondern längs einer dazu parallelen Geraden $z = x + iy$ ($y = \text{const}$) genommen wird. Somit gilt

$$g(\lambda) = \int_{-\infty}^{\infty} e^{-a(x+iy)^2} e^{-i\lambda(x+iy)}\, dx$$

$$= e^{ay^2+\lambda y} \int_{-\infty}^{\infty} e^{-ax^2-2aixy-i\lambda x}\, dx$$

$$= e^{ay^2+\lambda y} \int_{-\infty}^{\infty} e^{-ax^2-ix(2ay+\lambda)}\, dx.$$

Die Konstante y wählen wir jetzt so, daß der Imaginärteil im Exponenten des Integranden entfällt, d. h. $y = -\lambda/2a$. Dann ist

$$g(\lambda) = e^{a(\lambda^2/4a^2)-(\lambda^2/2a)} \int_{-\infty}^{\infty} e^{-ax^2} dx = e^{-\lambda^2/4a} \sqrt{\frac{\pi}{a}}$$

wegen

$$\int_{-\infty}^{\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}}.$$

Insbesondere erhalten wir für $a = 1/2$

$$f(x) = e^{-x^2/2}, \qquad g(\lambda) = \sqrt{2\pi}\, e^{-\lambda^2/2},$$

d. h., die Fouriertransformierte $g(\lambda)$ der Funktion $e^{-x^2/2}$ ist (bis auf einen konstanten Faktor) wieder dieselbe Funktion.

8.4.2. Die Grundeigenschaften der Fouriertransformation. Aus der Definitionsformel der Fouriertransformierten (1) ergeben sich eine Reihe von Eigenschaften, die wir jetzt untersuchen wollen. Zur Abkürzung werden wir die Fouriertransformierte der Funktion f mit dem Symbol $F[f]$ bezeichnen. Mit anderen Worten bezeichnen wir mit F den linearen Operator auf $L_1(-\infty, \infty)$, der jeder Funktion dieses Raumes ihre Fouriertransformierte[1]) zuordnet. Diesen Operator nennt man *Fouriertransformation.*

1. *Konvergiert die Folge* $\{f_n\}$ *von Funktionen aus* $L_1(-\infty, \infty)$ *in der Metrik des Raumes* $L_1(-\infty, \infty)$, *so konvergiert die Folge ihrer Fouriertransformierten* $g_n = F[f_n]$ *gleichmäßig auf der ganzen reellen Achse.*

[1]) Die Fouriertransformierte gehört im allgemeinen nicht zu L_1.

Diese Behauptung ergibt sich unmittelbar aus der trivialen Abschätzung

$$|g_n(\lambda) - g_m(\lambda)| \leqq \int_{-\infty}^{\infty} |f_n(x) - f_m(x)|\, dx.$$

2. *Die Fouriertransformierte g einer (absolut) integrierbaren Funktion f ist eine beschränkte stetige Funktion, die für $|\lambda| \to \infty$ gegen Null strebt.*

Die Beschränktheit der Funktion $g = F[f]$ ist unmittelbar aus der Abschätzung

$$|g(\lambda)| \leqq \int_{-\infty}^{\infty} |f(x)|\, dx$$

ersichtlich. Ferner sei f zunächst die charakteristische Funktion eines Intervalls (a, b). Dann ist die Fouriertransformierte

$$g(\lambda) = \int_a^b e^{-i\lambda x}\, dx = \frac{e^{i\lambda a} - e^{i\lambda b}}{i\lambda}$$

offenbar stetig und strebt für $|\lambda| \to \infty$ gegen Null. Wegen der Linearität der Fouriertransformation F ist also auch die Fouriertransformierte jeder Treppenfunktion (d. h. jeder Linearkombination von charakteristischen Funktionen von Intervallen) stetig und strebt für $\lambda \to \pm\infty$ gegen Null. Ist schließlich f eine Funktion aus L_1, so existiert eine Folge von Treppenfunktionen, die in $L_1(-\infty, \infty)$ gegen f konvergiert, denn die Treppenfunktionen liegen dicht in $L_1(-\infty, \infty)$. Auf Grund von Eigenschaft 1 konvergiert die Folge der Funktionen $g_n = F[f_n]$ auf der ganzen reellen Achse gleichmäßig gegen $g = F[f]$. Dann ist aber auch die Grenzfunktion g stetig und strebt für $|\lambda| \to \infty$ gegen Null.

Aufgaben

1. Es ist zu beweisen, daß die Fouriertransformierte g einer integrierbaren Funktion f auf der ganzen reellen Achse gleichmäßig stetig ist.

2. B sei der Raum aller gleichmäßig stetigen Funktionen auf $(-\infty, \infty)$, die im Unendlichen gegen Null streben. Es ist zu zeigen, daß die Fouriertransformation F ein Operator von $L_1(-\infty, \infty)$ in B mit $\|F\| = 1$ und $\operatorname{Ker} f = 0$ ist.

3. *Ist f auf jedem endlichen Intervall absolut stetig und $f' \in L_1(-\infty, \infty)$, so gilt die Gleichheit*

$$F[f'] = i\lambda\, F[f].$$

Somit entspricht der Differentiation einer Funktion (unter den obigen Bedingungen) die Multiplikation ihrer Fouriertransformierten mit $i\lambda$. Um das zu zeigen, schreiben wir die absolut stetige Funktion f in der Form

$$f(x) = f(0) + \int_0^x f'(t)\, dt.$$

Wegen der Integrierbarkeit von f' besitzt hier der rechts stehende Ausdruck für $x \to \infty$ und $x \to -\infty$ einen Grenzwert. Dieser Grenzwert kann nur Null sein, weil die Funktion f sonst nicht auf der ganzen Geraden integrierbar wäre. Damit ergibt sich durch partielle Integration

$$F[f'](\lambda) = \int_{-\infty}^{\infty} f'(x)\, e^{-i\lambda x}\, dx$$

$$= f(x)\, e^{-i\lambda x}\Big|_{-\infty}^{\infty} + i\lambda \int_{-\infty}^{\infty} f(x)\, e^{-i\lambda x}\, dx$$

$$= i\lambda\, F[f](\lambda),$$

was zu beweisen war.

Mit Hilfe der gleichen Überlegungen erhält man die Beziehung

$$F[f^{(k)}] = (i\lambda)^k\, F[f], \tag{5}$$

wenn $f^{(k-1)}$ auf jedem endlichen Intervall absolut stetig ist und wenn $f, \ldots, f^{(k)}$ zu $L_1(-\infty, \infty)$ gehören.

4. *Der Zusammenhang zwischen dem Grad der Glattheit einer Funktion und der Geschwindigkeit des Fallens ihrer Fouriertransformierten im Unendlichen.* Wir dividieren Gleichung (5) durch $(i\lambda)^k$ und erinnern daran, daß die Fouriertransformierte im Unendlichen stets gegen Null strebt (Eigenschaft 2). Daraus ergibt sich

$$|F[f]| = \frac{|F[f^{(k)}]|}{|\lambda|^k} \to 0,$$

wenn $f^{(k)}$ *integrierbar* ist. Unter den obigen Bedingungen fällt $F[f]$ im Unendlichen somit schneller als $1/|\lambda|^k$. Je größer also die Zahl der (zu L_1 gehörenden) Ableitungen von f ist, desto schneller fällt die Fouriertransformierte von f im Unendlichen.

5. *Wenn f'' existiert und zu $L_1(-\infty, \infty)$ gehört, ist $F[f]$ integrierbar.*

Unter den angegebenen Bedingungen ist $F[f]$ nämlich beschränkt und fällt im Unendlichen schneller als $1/\lambda^2$. Hieraus folgt die Integrierbarkeit.

Oben (Eigenschaft 4) haben wir gezeigt, daß die Fouriertransformierte einer Funktion f im Unendlichen um so schneller fällt, je mehr Ableitungen f besitzt. Es gilt auch eine dazu duale Aussage. Und zwar ist die Fouriertransformierte von f um so glatter, je schneller f fällt. Es gilt also folgende Aussage:

6. *Es sei sowohl die Funktion $f(x)$ als auch die Funktion $xf(x)$ integrierbar. Dann ist die Funktion $g = F[f]$ differenzierbar, und es gilt*

$$g'(\lambda) = F[-ix\, f(x)]. \tag{6}$$

Aus der Definitionsgleichung

$$g(\lambda) = \int_{-\infty}^{\infty} f(x)\, e^{-i\lambda x}\, dx$$

ergibt sich die Differenzierbarkeit von g, da das Integral nach λ differenzierbar ist. Dabei darf man unter dem Integral differenzieren, weil das Integral

$$-i \int_{-\infty}^{\infty} x\, f(x)\, e^{-i\lambda x}\, dx$$

(wegen der Integrierbarkeit von $xf(x)$) gleichmäßig bezüglich λ konvergiert. Folglich gilt Gleichung (6).

Gehören $f(x), xf(x), \ldots, x^p f(x)$ zu $L_1(-\infty, \infty)$, so zeigt man ganz analog die p-malige Differenzierbarkeit der Fouriertransformierten $g = F[f]$ und die Formeln

$$g^k(\lambda) = F[(-ik)^k f(x)] \qquad (k = 0, 1, \ldots, p).$$

7. Wenn man voraussetzt, daß die Funktion f im Unendlichen noch schneller fällt, wird die Fouriertransformierte g noch glatter sein. Gehört etwa $x^p f(x)$ für alle p zu $L_1(-\infty, \infty)$, so ist g unendlich oft differenzierbar. Wir nehmen jetzt an, daß $e^{\delta|x|} f(x) \in L_1(-\infty, \infty)$ ist für ein $\delta > 0$. Dann läßt sich $g(\lambda)$ von der reellen Achse analytisch auf einen Streifen in der komplexen Ebene $\zeta = \lambda + i\mu$ fortsetzen. Dabei ist die Breite des Streifens um so größer, je größer δ ist. In jedem Fall läßt sich g mindestens auf den Streifen $|\mu| < \delta$ analytisch fortsetzen. Denn das Integral

$$\int_{-\infty}^{\infty} f(x)\, e^{-ix\zeta}\, dx$$

konvergiert offenbar für $|\mu| < \delta$ und stellt eine stetige Funktion dar, die auf der reellen Achse mit der Fouriertransformierten von f übereinstimmt. Daß diese Funktion für $|\mu| < \delta$ im Sinne der analytischen Funktionen differenzierbar ist, wird genau wie Eigenschaft 6 bewiesen.

8.4.3. Die Vollständigkeit der Hermiteschen und der Laguerreschen Funktionen. Mit Hilfe der Überlegungen des vorhergehenden Absatzes kann man folgendes zeigen: *Ist die meßbare Funktion f fast überall auf dem Intervall (a, b), $-\infty \leqq a < b \leqq \infty$, von Null verschieden und genügt sie der Bedingung $|f(x)| \leqq Ce^{-\delta|x|}$ mit einem $\delta > 0$, so ist das System der Funktionen $\{x^n f(x)\}$, $n = 0, 1, 2, \ldots$, in $L_2(a, b)$ vollständig.*

Hieraus folgt insbesondere, daß die Hermiteschen Funktionen in $L_2(-\infty, \infty)$ und die Laguerreschen Funktionen in $L_2(0, \infty)$ ein vollständiges System bilden (siehe 7.3.8. und 7.3.9.).

Wir wollen die Behauptung über die Vollständigkeit beweisen. Dazu nehmen wir an, daß das System $\{x^n f(x)\}$ nicht vollständig ist. Dann gibt es auf Grund des Satzes von HAHN-BANACH eine von Null verschiedene Funktion $h \in L_2(-\infty, \infty)$ mit

$$\int_{-\infty}^{\infty} x^n f(x)\, h(x)\, dx = 0 \qquad (n = 0, 1, 2, \ldots).$$

(Wir haben hier den Satz über die allgemeine Form eines stetigen linearen Funktionals im Hilbertraum ausgenutzt. Betrachtet man den komplexen $L_2(a, b)$, so muß man $\overline{h(x)}$ an Stelle von $h(x)$ schreiben.) Offenbar ist $fh \in L_1(a, b)$, darüber hinaus ist sogar $e^{\delta_1|x|} fh \in L_1(a, b)$ für jedes $\delta_1 < \delta$. Im weiteren wollen wir annehmen, daß die Funktionen $f(x)$ und $h(x)$ auf der ganzen reellen Achse definiert sind, indem wir sie nötigenfalls außerhalb von (a, b) mit Null fortsetzen. Es sei g die Fouriertransformierte der Funktion fh, d. h.

$$g(\lambda) = \int_{-\infty}^{\infty} f(x)\, h(x)\, e^{-i\lambda x}\, dx.$$

Aus dem oben Gesagten folgt, daß sich die Funktion g analytisch auf den Streifen $|\mathrm{Im}\, \zeta| < \delta$ fortsetzen läßt. Dabei verschwinden alle Ableitungen dieser Funktion im Nullpunkt, so daß $g(\lambda) \equiv 0$ ist. Wegen der in 8.4.1. gezeigten Eindeutigkeit der Fouriertransformation ist daher $f(x)\, h(x) = 0$ fast überall. Folglich ist $h(x) = 0$ fast überall in (a, b), da $f(x)$ dort fast überall von Null verschieden war. Das widerspricht aber der Annahme, daß h eine von Null verschiedene Funktion ist. Dieser Widerspruch beweist die Vollständigkeit des Systems $\{x^n f(x)\}$.

8.4.4. Die Fouriertransformation der schnell fallenden unendlich oft differenzierbaren Funktionen. Beim Übergang von einer Funktion f zu ihrer Fouriertransformierten g vertauschen die Eigenschaften der Glattheit und der Geschwindigkeit des Fallens im Unendlichen ihre Rollen. Wenn man das berücksichtigt, kann man leicht Klassen von Funktionen angeben, die bezüglich der Fouriertransformation invariant sind.

Es sei S_∞ die Gesamtheit aller unendlich oft differenzierbaren Funktionen auf der reellen Achse, für die (von der jeweiligen Funktion f und den Zahlen p und q abhängige) Konstanten C_{pq} mit

$$|x^p f^{(q)}(x)| \leqq C_{pq} \tag{7}$$

existieren. Wir werden zeigen, daß mit $f \in S_\infty$ auch $g = F[f] \in S_\infty$ ist.

Zunächst ergibt sich aus (7) die absolute Integrierbarkeit aller Funktionen $x^p f^{(q)}(x)$. Denn wegen (7) gilt für alle p und q

$$|x^p f^{(q)}(x)| \leqq C_{p+2,q}/x^2,$$

d. h., die Funktion $x^p f^{(q)}(x)$ fällt nicht langsamer als $1/x^2$. Hieraus folgt aber, daß die Funktion $F[f]$ Ableitungen aller Ordnungen besitzt. Schließlich folgt gemäß 8.4.2.

aus der Integrierbarkeit von $f^{(q)}(x)$, $q = 1, 2, \ldots$, daß $g = F[f]$ im Unendlichen schneller als $1/|\lambda|^q$ fällt. Dann ist aber jede der Funktionen

$$(i\lambda)^q g^{(p)}(\lambda)' = (-i)^q\, F[(x^p f(x))^{(q)}]$$

als Fouriertransformierte einer integrierbaren Funktion durch eine Konstante D_{pq} beschränkt. Also gehört $g = F[f]$ für jede Funktion $f \in S_\infty$ ebenfalls zu S_∞.

Umgekehrt sei $g \in S_\infty$. Nach dem oben Gezeigten liegt dann die Funktion

$$f^*(x) = \int_{-\infty}^{\infty} g(\lambda)\, e^{-i\lambda x}\, dx$$

in S_∞. Setzen wir $f(x) = f^*(-x)/2\pi$, so ist offenbar auch $f \in S_\infty$. Zugleich folgt nach der Umkehrformel

$$g(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f^*(x)\, e^{i\lambda x} dx = \int_{-\infty}^{\infty} f(x)\, e^{-i\lambda x} dx,$$

d. h., g ist die Fouriertransformierte der Funktion $f \in S_\infty$. Somit führt die Fouriertransformation die Klasse S_∞ wieder in ganz S_∞ über. Es ist klar, daß diese Abbildung auch umkehrbar eindeutig ist.

Aufgabe. Es sei $f \in S_\infty$ und $\int_{-\infty}^{\infty} x^p f(x)\, dx = 0$ für alle $p \geqq 0$. Folgt hieraus $f(x) \equiv 0$?

8.4.5. Fouriertransformation und Faltung. Die Funktionen f_1 und f_2 seien auf der ganzen reellen Achse integrierbar. Die Funktion

$$f(x) = \int_{-\infty}^{\infty} f_1(\xi)\, f_2(x - \xi)\, d\xi$$

heißt dann *Faltung* von f_1 und f_2. Die Funktion $f(x)$ ist für fast alle x definiert und integrierbar, denn das zweifache Integral

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_1(\xi)\, f_2(x - \xi)\, d\xi\, dx$$

existiert wegen

$$\int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} |f_1(\xi)\, f(\eta)|\, d\xi \right) d\eta < \infty$$

(vgl. die Bemerkung zum Satz von Fubini in 5.6.4.). Folglich existiert auch das Integral

$$\int_{-\infty}^{\infty} f(x)\, dx = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} f_1(\xi)\, f_2(x - \xi)\, d\xi.$$

Die Funktion f wird mit dem Symbol $f_1 * f_2$ bezeichnet.

Wir wollen die Fouriertransformierte der Faltung von zwei Funktionen aus L_1 berechnen. Wenn wir den Satz von FUBINI anwenden und $x - \xi = \eta$ setzen, erhalten wir

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)\, e^{-i\lambda x} dx &= \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} f_1(\xi)\, f_2(x - \xi)\, d\xi \right\} e^{-i\lambda x} dx \\ &= \int_{-\infty}^{\infty} f_1(\xi) \left\{ \int_{-\infty}^{\infty} f_2(x - \xi)\, e^{-i\lambda x} dx \right\} d\xi \\ &= \int_{-\infty}^{\infty} f_1(\xi) \left\{ \int_{-\infty}^{\infty} f_2(\eta)\, e^{-i\lambda\eta} e^{-i\lambda\xi} d\eta \right\} d\xi \\ &= \int_{-\infty}^{\infty} f_2(\eta)\, e^{-i\lambda\eta}\, d\eta \cdot \int_{-\infty}^{\infty} f_1(\xi)\, e^{-i\lambda\xi}\, d\xi.\end{aligned}$$

Somit gilt

$$F[f_1 * f_2] = F[f_1] \cdot F[f_2],$$

d. h., *die Fouriertransformation führt die Faltung in eine einfachere Operation, die Multiplikation von Funktionen, über.* Diese Tatsache spielt in vielen Anwendungen der Fouriertransformation eine wichtige Rolle.

8.4.6. Die Anwendung der Fouriertransformation zur Lösung der Wärmeleitungsgleichung. Die Anwendbarkeit der Fouriertransformation auf Differentialgleichungen beruht darauf (vgl. 8.4.2.), daß sie die Differentiation in eine Multiplikation mit der unabhängigen Veränderlichen überführt. Folglich geht eine lineare Differentialgleichung mit konstanten Koeffizienten

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1} y' + a_n y = \varphi(x) \tag{8}$$

bei der Fouriertransformation in die algebraische Gleichung

$$(i\lambda)^n z + a_1 (i\lambda)^{n-1} z + \cdots + a_{n-1} i\lambda z + a_n z = \psi(\lambda) \tag{9}$$

mit $z = F[y]$ und $\psi = F[\varphi]$ über. Für gewöhnliche Differentialgleichungen eröffnet dieser Zugang jedoch keine wesentlich neuen Perspektiven, denn die Lösung linearer Gleichungen mit konstanten Koeffizienten bereitet auch ohne die Fouriertransformation keine Mühe. Außerdem ist der Übergang von (8) zu (9) nur möglich, wenn die gesuchte Funktion $y = y(x)$ auf der ganzen reellen Achse integrierbar ist. Für Lösungen linearer Differentialgleichungen mit konstanten Koeffizienten trifft das aber allgemein gar nicht zu.

Wesentlich wichtiger ist die Anwendung der Fouriertransformation auf partielle Differentialgleichungen. Sie gestattet nämlich unter bestimmten Bedingungen, die Lösung einer solchen Gleichung auf die Lösung einer gewöhnlichen Differential-

gleichung zurückzuführen. Wir wollen das am Beispiel des Cauchyproblems für die Wärmeleitungsgleichung demonstrieren.

Wir suchen eine Lösung der Gleichung

$$\frac{\partial u(x, t)}{\partial t} = \frac{\partial^2 u(x, t)}{\partial x^2} \tag{10}$$

für $-\infty < x < \infty$ und $t \geqq 0$, die für $t = 0$ den Anfangswert $u_0(x)$ annimmt. Der physikalische Sinn dieser Aufgabe besteht darin, die Temperaturverteilung in einem unendlichen wärmeleitenden Stab zu jedem Zeitpunkt $t > 0$ zu bestimmen, wenn die Temperaturverteilung im Anfangsmoment $t = 0$ gleich $u_0(x)$ ist.

Wir setzen voraus, daß $u_0(x)$, $u_0'(x)$ und $u_0''(x)$ zu $L_1(-\infty, \infty)$ gehören, und suchen eine Lösung $u(x, t)$ der Gleichung (10) mit folgenden Eigenschaften:

1. Die Funktionen $u(x, t)$, $u_x(x, t)$ und $u_{xx}(x, t)$ sind für jedes feste $t \geqq 0$ auf der ganzen x-Achse integrierbar.

2. Die Funktion $u_t(x, t)$ besitzt in jedem endlichen Intervall $0 \leqq t \leqq T$ eine (von t unabhängige) integrierbare Majorante $f(x)$:

$$|u_t(x, t)| \leqq f(x), \quad \int_{-\infty}^{\infty} f(x)\, dx < \infty.$$

Wenden wir auf Gleichung (10) die Fouriertransformation bezüglich x an, so erhalten wir rechts

$$F[u_{xx}(x, t)] = -\lambda^2 v(\lambda, t) \qquad \text{mit } v(\lambda, t) = F[u(x, t)]$$

und links unter Ausnutzung von Eigenschaft 2

$$F[u_t] = \int_{-\infty}^{\infty} u_t(x, t)\, e^{-i\lambda x} dx = \frac{\partial}{\partial t} \int_{-\infty}^{\infty} u(x, t)\, e^{-i\lambda x} dx = v_t(\lambda, t).$$

Somit führt die Fouriertransformation die Gleichung (10) in die gewöhnliche Differentialgleichung

$$v_t(\lambda, t) = -\lambda^2 v\,(\lambda, t)$$

über, für die nun eine Lösung $v(\lambda, t)$ mit

$$v(\lambda, 0) = v_0(\lambda) = F[u_0(x)] = \int_{-\infty}^{\infty} u_0(x)\, e^{-i\lambda x}\, dx$$

gesucht wird. Eine solche Lösung ist offenbar

$$v(\lambda, t) = e^{-\lambda^2 t}\, v_0(\lambda).$$

Um jetzt eine Lösung der ursprünglichen Aufgabe zu erhalten, muß man eine Funktion $u(x, t)$ finden, deren Fouriertransformierte die Funktion $v(\lambda, t)$ ist. Nach 8.4.1.,

Beispiel 4, gilt

$$e^{-\lambda^2 t} = F\left[\frac{1}{2\sqrt{\pi t}}\, e^{-x^2/4t}\right].$$

Daher ist

$$v(\lambda, t) = F\left[\frac{1}{2\sqrt{\pi t}}\, e^{-x^2/4t}\right] \cdot F[u_0(x)] = F\left[\frac{1}{2\sqrt{\pi t}}\, e^{-x^2/4t} * u_0(x)\right],$$

d. h.

$$u(x, t) = \frac{1}{2\sqrt{\pi t}} \int\limits_{-\infty}^{\infty} e^{-\xi^2/4t} u_0(x - \xi)\, d\xi .$$

Damit haben wir das sogenannte *Poissonsche Integral* als Lösung der Wärmeleitungsgleichung erhalten.

8.4.7. Die Fouriertransformation von Funktionen mehrerer Veränderlicher. Die Fouriertransformation, die bisher nur für Funktionen einer Veränderlichen betrachtet wurde, läßt sich auch leicht für Funktionen mehrerer Veränderlicher erklären.

Ist $f(x_1, x_2, \ldots, x_n)$ eine auf dem ganzen n-dimensionalen Raum $\boldsymbol{R}^n$ integrierbare Funktion, so heißt die Funktion

$$g(\lambda_1, \lambda_2, \ldots, \lambda_n) = \int\limits_{-\infty}^{\infty} \cdots \int\limits_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_n)\, e^{-i(x_1\lambda_1 + x_2\lambda_2 + \cdots + x_n\lambda_n)}\, dx_1 \cdots dx_n$$

Fouriertransformierte von f.

Dieses n-fache Integral existiert offenbar, da $f(x_1, x_2, \ldots, x_n)$ integrierbar ist. Nach dem Satz von Fubini kann man es auch in Form eines iterierten Integrals schreiben:

$$g(\lambda_1, \lambda_2, \ldots, \lambda_n) = \int\limits_{-\infty}^{\infty} \left\{ \cdots \int\limits_{-\infty}^{\infty} \left\{ \int\limits_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_n)\, e^{-ix_1\lambda_1}\, dx_1 \right\} \times e^{-ix_2\lambda_2} dx_2 \cdots \right\} e^{-ix_n\lambda_n}\, dx_n . \tag{11}$$

Anders ausgedrückt, kann man die Fouriertransformierte einer Funktion von n Veränderlichen erhalten, indem man die eindimensionale Fouriertransformation (in beliebiger Reihenfolge) sukzessiv auf jede Veränderliche einzeln anwendet. Wenn wir nacheinander jede der n Operationen auf der rechten Seite von Gleichung (11)

umkehren, so erhalten wir die Formel

$$f(x_1, x_2, \ldots, x_n) = \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} \left\{ \cdots \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} g(\lambda_1, \lambda_2, \ldots, \lambda_n)\, e^{i x_n \lambda_n}\, d\lambda_n \right] \times e^{i x_{n-1} \lambda_{n-1}}\, d\lambda_{n-1} \cdots \right\} e^{i x_1 \lambda_1}\, d\lambda_1 .$$

Die rechte Seite kann man hier auch als n-faches Integral schreiben:

$$f(x_1, x_2, \ldots, x_n) = \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(\lambda_1, \lambda_2, \ldots, \lambda_n)\, e^{i(x_1\lambda_1 + x_2\lambda_2 + \cdots + x_n\lambda_n)}\, d\lambda_1 \cdots d\lambda_n . \tag{12}$$

Hierbei ist jedoch die Funktion $g(\lambda_1, \lambda_2, \ldots, \lambda_n)$ im allgemeinen nicht notwendig im ganzen $\boldsymbol{R}^n$ integrierbar. Daher muß man jeweils angeben, in welchem Sinne dieses Integral zu verstehen ist und unter welchen Bedingungen an $f(x_1, x_2, \ldots, x_n)$ die Gleichung (12) gilt.

Eine der möglichen Antworten auf diese Fragen liefert der folgende Satz.

Satz 1. *Die Funktion $f(x_1, x_2, \ldots, x_n)$ sei im ganzen $\boldsymbol{R}^n$ integrierbar und genüge den Bedingungen*

$$\left.\begin{aligned} &|f(x_1 + t_1, x_2, \ldots, x_n) - f(x_1, x_2, \ldots, x_n)| \leqq C\, |t_1|^a, \\ &|f(x_1, x_2 + t_2, \ldots, x_n) - f(x_1, x_2, \ldots, x_n)| \leqq C(x_1)\, |t_2|^a, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &|f(x_1, x_2, \ldots, x_n + t_n) - f(x_1, x_2, \ldots, x_n) \leqq C(x_1, x_2, \ldots, x_{n-1})\, |t_n|^a \\ &\text{mit } 0 \leqq a \leqq 1 \text{ und } \int_{-\infty}^{\infty} C(x_1)\, dx_1 < \infty, \ldots, \\ &\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} C(x_1, x_2, \ldots, x_{n-1})\, dx_1\, dx_2 \cdots dx_{n-1} < \infty . \end{aligned}\right\} \tag{13}$$

Dann ist die Umkehrformel (12) *gültig, wenn man dabei das Integral in folgendem Sinne versteht:*

$$\frac{1}{(2\pi)^n} \lim_{N_1 \to \infty} \int_{-N_1}^{N_1} \left\{ \cdots \lim_{N_{n-1} \to \infty} \int_{-N_{n-1}}^{N_{n-1}} \left[\lim_{N_n \to \infty} \int_{-N_n}^{N_n} g(\lambda_1, \lambda_2, \ldots, \lambda_n)\, e^{i x_n \lambda_n}\, d\lambda_n \right] \times e^{i x_{n-1} \lambda_{n-1}}\, d\lambda_{n-1} \cdots \right\} e^{i x_1 \lambda_1}\, d\lambda_1 .$$

Beweis. Wegen der Integrierbarkeit von $f(x_1, x_2, \ldots, x_n)$ im $\boldsymbol{R}^n$ ist diese Funktion nach dem Satz von FUBINI für fast alle $x_2, \ldots, x_n$ bezüglich x_1 integrierbar. Folglich existiert die Funktion

$$f_1(\lambda_1, x_2, \ldots, x_n) = \int_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_n)\, e^{-ix_1\lambda_1}\, dx_1 .$$

Nach (13) erfüllt die Funktion $f(x_1, x_2, \ldots, x_n)$ als Funktion von x_1 die Bedingungen von Satz 1 aus 8.3. Daher kann man $f(x_1, x_2, \ldots, x_n)$ nach der Umkehrformel durch f_1 ausdrücken:

$$f(x_1, x_2, \ldots, x_n) = \lim_{N_1\to\infty} \frac{1}{2\pi} \int_{-N_1}^{N_1} f_1(\lambda_1, x_2, \ldots, x_n)\, e^{ix_1\lambda_1}\, d\lambda_1 .$$

Aus der Gleichung

$$f_2(\lambda_1, \lambda_2, x_3, \ldots, x_n) = \int_{-\infty}^{\infty} f_1(\lambda_1, x_2, \ldots, x_n)\, e^{-ix_2\lambda_2}\, dx_2$$

folgt, wieder nach (13), die Gültigkeit der Umkehrformel

$$f_1(\lambda_1, x_2, \ldots, x_n) = \lim_{N_2\to\infty} \frac{1}{2\pi} \int_{-N_2}^{N_2} f_2(\lambda_1, \lambda_2, \ldots, x_n)\, e^{ix_2\lambda_2}\, d\lambda_2 .$$

Damit gilt also

$$f(x_1, x_2, \ldots, x_n) = \lim_{N_1\to\infty} \frac{1}{2\pi} \int_{-N_1}^{N_1} \left\{ \lim_{N_2\to\infty} \int_{-N_2}^{N_2} f_2(\lambda_1, \lambda_2, \ldots, x_n)\, e^{ix_2\lambda_2}\, d\lambda_2 \right\} e^{ix_1\lambda_1}\, d\lambda_1 .$$

Wenn wir $f_3(\lambda_1, \lambda_2, \lambda_3, \ldots, x_n)$ usw. in analoger Weise definieren, erhalten wir genau Formel (12).

Die Fouriertransformation von Funktionen mehrerer Veränderlicher findet in der Theorie der partiellen Differentialgleichungen breite Anwendung. Wir betrachten beispielsweise die Gleichung

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}, \tag{14}$$

die den Prozeß der Wärmeausbreitung in der Ebene beschreibt. Zum Zeitpunkt $t = 0$ sei die Temperatur vorgegeben:

$$u(0, x, y) = u_0(x, y) .$$

Auf Gleichung (14) können wir die Fouriertransformation bezüglich der Veränderlichen x und y anwenden, wenn wir über die gesuchte Lösung dieser Gleichung analoge Bedingungen voraussetzen, wie sie in 8.4.6. angegeben wurden. Als Resultat erhalten wir dann die gewöhnliche Differentialgleichung

$$\frac{dv}{dt} = -(\lambda^2 + \sigma^2)\, v \tag{15}$$

mit

$$v(t, \lambda, \sigma) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u(t, x, y)\, e^{-i(\lambda x + \sigma y)}\, dx\, dy.$$

Nachdem man Gleichung (15) gelöst hat, findet man mit Hilfe der Umkehrformel eine Lösung der Ausgangsgleichung (14).

8.5. Die Fouriertransformation im Raum $L_2(-\infty, \infty)$

8.5.1. Der Satz von Plancherel. Wir kehren zunächst noch einmal zu den Resultaten zurück, die wir über Fourierreihen erhalten haben. Um eine größere Analogie zur Fouriertransformation zu erhalten, wollen wir die Fourierreihen in komplexer Form betrachten. Wir nehmen also das vollständige orthogonale System der Funktionen e^{inx} $(n = 0, \pm 1, \pm 2, \ldots)$ auf dem Intervall $[-\pi, \pi]$ und ordnen jeder auf diesem Intervall integrierbaren Funktion f die Folge ihrer Fourierkoeffizienten c_n zu,

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\, e^{-inx}\, dx \qquad (n = 0, \pm 1, \pm 2, \ldots).$$

Ist die Funktion f nicht nur integrierbar, sondern sogar quadratisch integrierbar, so genügen die Fourierkoeffizienten von f der Bedingung

$$\sum_{n=-\infty}^{\infty} |c_n|^2 < \infty.$$

Umgekehrt ist auch jede solche Folge die Folge der Fourierkoeffizienten einer quadratisch integrierbaren Funktion. Also ist der Übergang von einer quadratisch integrierbaren Funktion zur Gesamtheit ihrer Fourierkoeffizienten eine Abbildung des Hilbertraumes L_2 auf den Hilbertraum l_2. Diese Abbildung ist linear und genügt der Parsevalschen Gleichung

$$2\pi \sum_{n=-\infty}^{\infty} |c_n|^2 = \int_{-\pi}^{\pi} |f(x)|^2\, dx \tag{1}$$

(d. h., diese Abbildung ist bis auf einen Zahlenfaktor isometrisch).

Jetzt wenden wir uns wieder der Fouriertransformation von Funktionen auf der ganzen reellen Achse zu und wollen sehen, ob diese Abbildung auch als linearer Operator im komplexen Raum $L_2(-\infty, \infty)$ aufgefaßt werden kann. Die Hauptschwierigkeit besteht hier darin, daß eine quadratisch integrierbare Funktion auf der reellen Achse nicht notwendig zu $L_1(-\infty, \infty)$ gehört. Die Fouriertransformation braucht also gar nicht in dem in 8.4. erklärten Sinne zu existieren. Für jedes $f \in L_2(-\infty, \infty)$ kann man jedoch die Fouriertransformation in einem etwas anderen Sinne definieren. Dabei erhält man folgenden Satz, den man als Analogon zur Parsevalschen Gleichung (1) ansehen kann.

Satz (Plancherel, 1910). *Für jede Funktion $f \in L_2(-\infty, \infty)$ und beliebiges $N > 0$ stellt das Integral*

$$g_N(\lambda) = \int_{-N}^{N} f(x)\, e^{-i\lambda x}\, dx$$

eine Funktion aus $L_2(-\infty, \infty)$ dar. Für $N \to \infty$ konvergieren die Funktionen g_N in der Metrik des Raumes L_2 gegen eine Funktion g mit

$$\int_{-\infty}^{\infty} |g(\lambda)|^2\, d\lambda = 2\pi \int_{-\infty}^{\infty} |f(x)|^2\, dx. \tag{2}$$

(Diese Funktion g nennt man die *Fouriertransformierte* der Funktion $f \in L_2$.) *Gehört f auch noch zu $L_1(-\infty, \infty)$, so stimmt die zugehörige Funktion g mit der Fouriertransformierten von f im üblichen Sinne überein.*

Beweis. Die Grundidee des Beweises besteht darin, daß man die Gültigkeit von (2) zunächst für Funktionen aus der Klasse S_∞ aller unendlich oft differenzierbaren schnell fallenden Funktionen nachweist. Diese Klasse liegt dicht in $L_2(-\infty, \infty)$, und daher kann man Gleichung (2) durch Stetigkeitsbetrachtungen auf ganz $L_2(-\infty, \infty)$ ausdehnen. Diese Idee wollen wir jetzt im einzelnen realisieren.

1. Es seien $f_1, f_2 \in S_\infty$ und g_1 bzw. g_2 die zugehörigen Fouriertransformierten. Dann gilt

$$\int_{-\infty}^{\infty} f_1(x)\, \overline{f_2(x)}\, dx = \int_{-\infty}^{\infty} \frac{1}{2\pi} \int_{-\infty}^{\infty} [g_1(\lambda)\, e^{i\lambda x}\, d\lambda]\, \overline{f_2(x)}\, dx$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[g_1(\lambda) \overline{\int_{-\infty}^{\infty} f_2(x)\, e^{-i\lambda x}\, dx} \right] d\lambda$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} g_1(\lambda)\, \overline{g_2(\lambda)}\, d\lambda.$$

Hier durfte die Integrationsreihenfolge vertauscht werden, da die Funktion

$$g_1(\lambda)\,\overline{f_2(x)}\,e^{i\lambda x}$$

in der (x, λ)-Ebene absolut integrierbar ist. Wenn wir in der erhaltenen Gleichung $f_1 = f_2 = f$ und $g_1 = g_2 = g$ setzen, ergibt sich die Gültigkeit von Formel (2) für beliebige Funktionen $f \in S_\infty$.

2. Jetzt sei f eine beliebige Funktion aus $L_2(-\infty, \infty)$, die außerhalb eines Intervalls $(-a, a)$ verschwindet. Dann ist f auf dem Intervall $(-a, a)$ integrierbar (d. h., f gehört zu $L_1(-a, a)$) und folglich auch auf der ganzen Geraden. Für diese Funktionen ist daher die Fouriertransformierte

$$g(\lambda) = \int_{-\infty}^{\infty} f(x)\,e^{-i\lambda x}\,dx$$

erklärt. Es sei nun $\{f_n\}$ eine Folge von Funktionen aus S_∞, die außerhalb von $(-a, a)$ verschwinden und die in der Norm des Raumes $L_2(-\infty, \infty)$ gegen f konvergieren. Da f und alle f_n nur auf einem endlichen Intervall verschieden sind, strebt die Folge $\{f_n\}$ auch in der Norm des Raumes $L_1(-\infty, \infty)$ gegen f. Also konvergiert die Folge $\{g_n\}$ auf der ganzen Geraden gleichmäßig gegen g (vgl. 8.4.2.). Außerdem ist $\{g_n\}$ Fundamentalfolge in $L_2(-\infty, \infty)$. Das ergibt sich auf Grund der Beziehung

$$\int_{-\infty}^{\infty} |g_n(\lambda) - g_m(\lambda)|^2\,d\lambda = 2\pi \int_{-\infty}^{\infty} |f_n(x) - f_m(x)|^2\,dx,$$

die wir schon für Funktionen aus S_∞ bewiesen haben. Somit konvergiert die Folge $\{g_n\}$ im Raum L_2, und zwar gegen dieselbe Funktion, gegen die sie auch gleichmäßig konvergiert. Daher kann man in der Gleichung

$$\|f_n\|^2 = \frac{1}{2\pi}\,\|g_n\|^2$$

zum Grenzwert für $n \to \infty$ übergehen. Damit erhalten wir die Gültigkeit von (2) für jede Funktion $f \in L_2$, die außerhalb eines Intervalls verschwindet.

3. Ist f schließlich eine beliebige Funktion aus L_2, so setzen wir

$$f_N(x) = \begin{cases} f(x) & \text{für } |x| \leqq N, \\ 0 & \text{für } |x| > N. \end{cases}$$

Offenbar gilt dann

$$\|f - f_N\| \to 0 \qquad \text{für } N \to \infty.$$

Die Funktion f_N gehört zu $L_1(-\infty, \infty)$, folglich existiert die Fouriertransformierte von f_N im gewöhnlichen Sinne,

$$g_N(\lambda) = \int_{-\infty}^{\infty} f_N(x)\,e^{-i\lambda x}\,dx = \int_{-N}^{N} f(x)\,e^{-i\lambda x}\,dx.$$

Nach dem zweiten Schritt ist

$$\|f_N - f_M\|^2 = \frac{1}{2\pi} \|g_N - g_M\|^2,$$

also konvergieren die Funktionen g_N in L_2 gegen einen Grenzwert g. Daher kann man in der Gleichung

$$\|f_N\|^2 = \frac{1}{2\pi} \|g_N\|^2$$

zum Grenzwert für $N \to \infty$ übergehen, woraus die Beziehung (2) für beliebige $f \in L_2(-\infty, \infty)$ folgt. Damit ist der erste Teil des Satzes von PLANCHEREL bewiesen.

Wenn jetzt die Funktion f sowohl zu $L_2(-\infty, \infty)$ als auch zu $L_1(-\infty, \infty)$ gehört, existiert die Fouriertransformierte

$$\tilde{g}(\lambda) = \int_{-\infty}^{\infty} f(x)\, e^{-i\lambda x}\, dx$$

im gewöhnlichen Sinne. Dabei konvergieren die Funktionen f_N in $L_1(-\infty, \infty)$ gegen f und die zugehörigen Fouriertransformierten g_N damit gleichmäßig gegen $\tilde{g}$. Außerdem haben wir aber bereits festgestellt, daß die Funktionen g_N in der Metrik des $L_2(-\infty, \infty)$ gegen einen Grenzwert g konvergieren. Hieraus folgt, daß $\tilde{g}$ und g übereinstimmen. Damit ist der Beweis beendet.

Folgerung. Aus (2) ergibt sich sofort, daß für beliebige $f_1, f_2 \in L_2(-\infty, \infty)$ die Gleichung

$$\int_{-\infty}^{\infty} f_1(x)\, \overline{f_2(x)}\, dx = \frac{1}{2\pi} \int_{\infty}^{\infty} g_1(\lambda)\, \overline{g_2(\lambda)}\, d\lambda$$

gilt. Zum Beweis braucht man (2) nur für die Funktion $f_1 + f_2$ aufzuschreiben und anschließend die Ausdrücke links und rechts zu vergleichen. Während Gleichung (2) besagte, daß die Fouriertransformation die Norm in L_2 unverändert läßt, bedeutet die letzte Gleichung, daß auch das Skalarprodukt erhalten bleibt.

8.5.2. Die Hermiteschen Funktionen. Der Satz von PLANCHEREL aus 8.5.1. besagt, daß man die Fouriertransformation als beschränkten linearen Operator F des Raumes $L_2(-\infty, \infty)$ auf sich ansehen kann.[1]) Wenn man in diesem Raum ein beliebiges

[1]) Anm. d. Übers.: Daß der Wertevorrat von F wieder ganz $L_2(-\infty, \infty)$ ist, sieht man wie folgt: Aus 8.4.4. ergibt sich, daß F die Menge S_∞ *auf* sich abbildet. Andererseits ist S_∞ dicht in $L_2(-\infty, \infty)$. Zusammen mit dem Satz von PLANCHEREL erhält man hieraus, daß das Bild von $L_2(-\infty, \infty)$ bei Anwendung von E mit ganz $L_2(-\infty, \infty)$ übereinstimmt.

orthonormiertes System auswählt, kann man den Operator F (wie auch jeden anderen linearen Operator) mit Hilfe einer unendlichen Matrix beschreiben. Die Gestalt dieser Matrix hängt dabei natürlich von der Wahl der Basis ab. Am einfachsten sieht die Matrix eines Operators in dem Fall aus, daß die entsprechende Basis nur aus Eigenfunktionen dieses Operators besteht. Die Matrix besitzt dann Diagonalform. Wir wollen sehen, ob eine solche Basis auch für die Fouriertransformation F existiert. Mit anderen Worten wollen wir feststellen, welche Funktionen aus $L_2(-\infty, \infty)$ Eigenfunktionen der Fouriertransformation F sind. Dazu bemerken wir, daß die Gleichung

$$\frac{d^2 f}{dx^2} - x^2 f = \mu f \tag{3}$$

durch die Fouriertransformation wieder in dieselbe Gleichung übergeführt wird[1]) (da die Operation d^2/dx^2 in die Multiplikation mit $-\lambda^2$ und die Multiplikation mit $-x^2$ in die Operation $d^2/d\lambda^2$ übergeht). Daher ist es naheliegend, Eigenfunktionen des Operators F als Lösungen der Gleichung (3) zu suchen. Wir wollen Lösungen dieser Gleichung suchen, die die Form

$$f = we^{-x^2/2}$$

besitzen, wobei w ein Polynom ist. Tragen wir diesen Ausdruck in (3) ein, so erhalten wir für w die Gleichung

$$w'' - 2xw' = (\mu + 1)\, w.$$

Wenn wir hier

$$w = a_0 + a_1 x + \cdots + a_n x^n \tag{4}$$

setzen, so gelangen wir zu

$$\begin{aligned}&\left(2a_2 + 3 \cdot 2a_3 x + \cdots + n(n-1)\, a_n x^{n-2}\right)\\ &\qquad - 2x\,(a_1 + 2a_2 x + \cdots + na_n x^{n-1})\\ &= (\mu + 1)\,(a_0 + a_1 x + \cdots + a_n x^n).\end{aligned}$$

Hieraus ergibt sich durch Vergleich der Koeffizienten gleicher Potenzen von x links und rechts

$$-2na_n = (\mu + 1)\, a_n, \qquad -2(n-1)\, a_{n-1} = (\mu + 1)\, a_{n-1}$$

und allgemein

$$k(k-1)\, a_k - 2(k-2)\, a_{k-2} = (\mu + 1)\, a_{k-2}. \tag{5}$$

[1]) Dabei wird natürlich vorausgesetzt, daß die gesuchte Funktion f entsprechend glatt ist und im Unendlichen hinreichend schnell fällt.

Wenn wir annehmen, daß der höchste Koeffizient a_n von Null verschieden ist, muß

$$\mu = -(2n+1) \quad \text{und} \quad a_{n-1} = 0$$

sein, d. h., μ muß eine negative ungerade ganze Zahl sein. Alle Koeffizienten des Polynoms w werden dann durch die Beziehung (5) bis auf einen konstanten Faktor bestimmt. Dabei sind die Koeffizienten $a_{n-1}, a_{n-3}, \ldots$ gleich Null. Die Koeffizienten $a_{n-2}, a_{n-4}, \ldots$ sind von Null verschieden; sie lassen sich nach der Rekursionsformel

$$a_{k-2} = \frac{k(k-1)}{2k-2n-4} a_k$$

berechnen (wenn a_n gegeben ist). Für w erhalten wir also die Formel

$$w_n(x) = a_n \left(x^n - \frac{n(n-1)}{4} x^{n-2} + \frac{n(n-1)(n-2)(n-3)}{4 \cdot 8} x^{n-4} - \cdots \right).$$

Damit haben wir ein System von Funktionen der Form

$$\varphi_n(x) = w_n(x)\, e^{-x^2/2} \qquad (n = 0, 1, 2, \ldots)$$

konstruiert. Jede dieser Funktionen gehört offenbar zu $L_2(-\infty, \infty)$ (wegen des Faktors $e^{-x^2/2}$). Darüber hinaus sind diese Funktionen paarweise orthogonal. Nach (3) gilt nämlich

$$\varphi_n''(x) - x^2 \varphi_n(x) = -(2n+1)\, \varphi_n(x),$$

$$\varphi_m''(x) - x^2 \varphi_m(x) = -(2m+1)\, \varphi_m(x).$$

Wenn wir die erste dieser Gleichungen mit φ_m und die zweite mit φ_n multiplizieren und anschließend die zweite Gleichung von der ersten subtrahieren, ergibt sich

$$\varphi_n''\varphi_m - \varphi_m''\varphi_n = 2(n-m)\, \varphi_n\varphi_m$$

bzw.

$$[\varphi_n'\varphi_m - \varphi_m'\varphi_n]' = 2(n-m)\, \varphi_n\varphi_m.$$

Durch Integration dieser Gleichung erhalten wir für $n \neq m$

$$\int_{-\infty}^{\infty} \varphi_n(x)\, \varphi_m(x)\, dx = \frac{1}{2(n-m)} \int_{-\infty}^{\infty} [\varphi_n'\varphi_m - \varphi_m'\varphi_n]'\, dx$$

$$= \frac{1}{2(n-m)} [\varphi_n'\varphi_m - \varphi_m'\varphi_n]_{-\infty}^{\infty} = 0.$$

Damit ist die Orthogonalität bewiesen.

Jedes der Elemente φ_n des so erhaltenen Orthogonalsystems ist ein Polynom n-ten Grades, multipliziert mit $e^{-x^2/2}$. Folglich müssen diese Elemente bis auf einen Zahlen-

faktor mit den Hermiteschen Funktionen übereinstimmen, die wir in 7.3. durch Orthogonalisierung der Folge

$$e^{-x^2/2},\ xe^{-x^2/2},\ \ldots,\ x^n e^{-x^2/2},\ \ldots$$

im Raum $L_2(-\infty, \infty)$ erhalten hatten.

Wir werden jetzt zeigen, daß die Funktionen $\{\varphi_n\}$ Eigenwerte der Fouriertransformation sind:

$$F\varphi_n = c_n\varphi_n. \tag{6}$$

Das ergibt sich aus den folgenden Fakten.

1. Gleichung (3) ist invariant bezüglich der Fouriertransformation F.

2. Gleichung (3) besitzt für jedes n, bis auf einen konstanten Faktor, nur eine Lösung der Form $P_n(x)\, e^{-x^2/2}$, wobei P_n ein Polynom n-ten Grades ist.

3. Die Fouriertransformation führt $x^n e^{-x^2/2}$ in $\left(i\,\frac{d}{dx}\right)^n e^{-x^2/2} = Q_n(x)\, e^{-x^2/2}$ über, wobei Q_n ein Polynom n-ten Grades ist (dies kann man leicht mittels Induktion beweisen).

Aus Gleichung (6) folgt für jedes ganze k die Beziehung

$$F^k\varphi_n = c_n{}^k\varphi_n.$$

Wendet man die Fouriertransformation viermal auf eine Funktion an, so wird diese Funktion (nach der Umkehrformel) nur mit dem Faktor $4\pi^2$ multipliziert. Daher ist

$$c_n{}^4 = 4\pi^2,$$

d. h., c_n kann nur die Werte $\pm\sqrt{2\pi}$ und $\pm i\sqrt{2\pi}$ annehmen.

Nimmt man also die Gesamtheit aller Hermiteschen Funktionen als Basis im Raum $L_2(-\infty, \infty)$, so wird die Fouriertransformation F durch eine Diagonalmatrix beschrieben, in deren Diagonale die Zahlen $\pm\sqrt{2\pi}$ und $\pm i\sqrt{2\pi}$ stehen.[1]

8.6. Die Laplacetransformation

8.6.1. Definition und grundlegende Eigenschaften der Laplacetransformation. Die Anwendbarkeit der Fouriertransformation auf Differentialgleichungen wird wesentlich dadurch eingeschränkt, daß diese Transformation nur für Funktionen erklärt

[1]) Wird die Fouriertransformation durch die Formel

$$F[f] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)\, e^{-i\lambda x}\, dx$$

(d. h. durch Formel (1′) aus 8.4. und nicht durch Formel (1)) definiert, so ist deren vierte Potenz der Einheitsoperator. Wenn man die Gesamtheit aller Hermiteschen Funktionen als Basis nimmt, erhält man für F eine Diagonalmatrix, in deren Hauptdiagonale die Zahlen ± 1 und $\pm i$ stehen.

ist, die auf der ganzen Achse integrierbar sind. Insbesondere existiert die Fouriertransformation nicht für Funktionen, die für $x \to -\infty$ oder $x \to +\infty$ wachsen. Solche Funktionen tauchen jedoch öfter bei der Lösung von Differentialgleichungen auf. Diese Schwierigkeiten kann man überwinden, wenn man die Fouriertransformation auf Distributionen ausdehnt. Über diesen Weg werden wir in 8.8. kurz berichten. Ein anderes mögliches Herangehen, das den Rahmen des klassischen Funktionsbegriffs und der klassischen Methoden der Analysis nicht verläßt, ist die Ersetzung der Fouriertransformation durch die sogenannte *Laplacetransformation.*

Die Funktion f (die im allgemeinen nicht auf der ganzen reellen Achse integrierbar zu sein braucht) sei nach Multiplikation mit $e^{-\gamma x}$ integrierbar. Dabei sei γ eine reelle Zahl. Dann ist das Integral

$$g(s) = \int_{-\infty}^{\infty} f(x)\, e^{-isx}\, dx = \int_{-\infty}^{\infty} f(x)\, e^{-ix\lambda} e^{x\mu}\, dx$$

für gewisse komplexe Zahlen $s = \lambda + i\mu$ konvergent, insbesondere auf der Geraden $\mu = -\gamma$. Auf dieser Geraden ist es die Fouriertransformierte der Funktion $f(x)\, e^{x\mu}$. Am wichtigsten für die Anwendung ist der Fall, daß die Funktion f den folgenden Bedingungen genügt:

$$\begin{aligned} |f(x)| &< C\, e^{\gamma_0 x} \quad \text{für } x \geqq 0, \\ f(x) &= 0 \qquad\quad \text{für } x < 0 \end{aligned} \tag{1}$$

(wobei γ_0 und C Konstanten sind). Die Voraussetzung über die Integrierbarkeit der Funktion $f(x)\, e^{-\gamma x}$ ist dabei für $\gamma > \gamma_0$ erfüllt. Das Integral

$$g(s) = \int_{-\infty}^{\infty} f(x)\, e^{-isx}\, dx = \int_{0}^{\infty} f(x)\, e^{-isx}\, dx \tag{2}$$

existiert also für alle $s = \lambda + i\mu$ mit $\mu < -\gamma_0$, d. h. in der Halbebene, die durch die Gerade $\operatorname{Im} s = -\gamma_0$ beschränkt ist. Dort stellt (2) die Fouriertransformierte der Funktion

$$f(x)\, e^{\mu x}$$

dar. Diese Funktion kann aus g mit Hilfe der Umkehrformel zurückerhalten werden (wir nehmen dabei an, daß diese Formel für f anwendbar ist),

$$f(x)\, e^{\mu x} = \frac{1}{2\pi} \int_{-\infty}^{\infty} g(s)\, e^{i\lambda x}\, d\lambda.$$

Daraus ergibt sich

$$f(x) = \frac{1}{2\pi} \int_{i\mu-\infty}^{i\mu+\infty} g(s)\, e^{isx}\, ds \quad (s = \lambda + i\mu). \tag{3}$$

Da die Funktion $f(x)\,e^{\mu x}$ für $\mu < -\gamma_0$ (nach (1)) exponentiell fällt, ist ihre Fouriertransformierte g und damit auch $g(s)\,e^{isx}$ in der Halbebene $\operatorname{Im} s < -\gamma_0$ analytisch.

Setzen wir in den Formeln (2) und (3) $p = is$ und bezeichnen $g(s)$ mit $\Phi(p)$, so erhalten wir

$$\Phi(p) = \int_0^\infty f(x)\,e^{-px}\,dx \tag{2'}$$

und

$$f(x) = \frac{1}{2\pi}\int_{-\mu-i\infty}^{-\mu+i\infty} \Phi(p)\,e^{px}\,\frac{dp}{i} = \frac{1}{2\pi i}\int_{-\mu-i\infty}^{-\mu+i\infty} \Phi(p)\,e^{px}\,dp. \tag{3'}$$

Die Funktion Φ ist dabei in der Halbebene $\operatorname{Re} p > \mu_0$ definiert und analytisch. Sie heißt *Laplacetransformierte* der Funktion f (mit den Eigenschaften (1)).

Die Laplacetransformation unterscheidet sich in ihren Eigenschaften nur wenig von der Fouriertransformation. Jedoch ist die Klasse der Funktionen, für die die Laplacetransformation erklärt ist, wesentlich verschieden von der Klasse $L_1(-\infty, \infty)$ der Funktionen, für die die Fouriertransformation existiert.

8.6.2. Die Anwendung der Laplacetransformation zur Lösung von Differentialgleichungen (Operatorenmethode). Die Laplacetransformation kann man zum Auffinden der Lösungen von Differentialgleichungen anwenden. Gegeben sei eine lineare Differentialgleichung mit konstanten Koeffizienten:

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_n y = b(x). \tag{4}$$

Gesucht ist eine Lösung dieser Gleichung mit den Anfangswerten

$$y(0) = y_0, \quad y'(0) = y_1, \quad \ldots, \quad y^{(n-1)}(0) = y_{n-1}. \tag{5}$$

Auf Gleichung (4) wenden wir die Laplacetransformation[1]) an, d. h., wir multiplizieren (4) mit e^{-px} und integrieren von 0 bis ∞. Es sei

$$Y(p) = \int_0^\infty y(x)\,e^{-px}\,dx$$

die Laplacetransformierte der Funktion y. Nach partieller Integration ergibt sich daraus die Laplacetransformierte der Ableitung y':

$$\int_0^\infty y'(x)\,e^{-px}\,dx = y(x)\,e^{-px}\Big|_0^\infty + p\int_0^\infty y(x)\,e^{-px}\,dx = pY(p) - y_0.$$

[1]) Die Anwendbarkeit der Laplacetransformation auf (4) ist leicht nachzuweisen, wenn $|b(x)|$ nicht zu schnell wächst.

Durch sukzessive Anwendung dieser Formel erhält man

$$\int_0^\infty y^{(n)}(x)\, e^{-px}\, dx = p(p^{n-1}Y(p) - y_{n-2} - py_{n-3} - \cdots - p^{n-2}y_0) - y_{n-1}$$
$$= p^n Y(p) - y_{n-1} - py_{n-2} - \cdots - p^{n-1}y_0$$
$$= p^n Y(p) - \sum_{k=0}^{n-1} p^{n-1-k} y_k .$$

Schließlich sei noch

$$B(p) = \int_0^\infty b(x)\, e^{-px}\, dx .$$

Im Ergebnis geht dann die Differentialgleichung (4) (mit den Anfangswerten (5)) bei der Laplacetransformation in die algebraische Gleichung

$$Q(p) + R(p)\, Y(p) = B(p)$$

über. Dabei ist B die Laplacetransformierte von b, Q ein Polynom $(n-1)$-ten Grades in p, das von den Koeffizienten der Gleichung und von den Anfangswerten abhängt, und schließlich ist

$$R(p) = \sum_{k=0}^{n} a_{n-k}\, p^k , \qquad a_0 = 1 ,$$

das charakteristische Polynom der Gleichung (4). Aus der erhaltenen Gleichung folgt

$$Y(p) = \frac{B(p) - Q(p)}{R(p)} ,$$

und hieraus erhält man nach der Umkehrformel die Lösung

$$y(x) = \frac{1}{2\pi i} \int_{-\mu - i\infty}^{-\mu + i\infty} \frac{B(p) - Q(p)}{R(p)}\, e^{px}\, dp .$$

Dieses Integral berechnet man gewöhnlich mit Hilfe des Residuensatzes.

Zur Lösung linearer Differentialgleichungen mit konstanten Koeffizienten wird auch die sogenannte *Operatorenmethode* verwendet. Sie besteht darin, daß man die linke Seite der Gleichung

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_n y = b(x)$$

als Resultat der Anwendung des Operators

$$A\left(\frac{d}{dx}\right) = \frac{d^n}{dx^n} + a_1 \frac{d^{n-1}}{dx^{n-1}} + \cdots + a_n \tag{6}$$

auffaßt. Eine Lösung der Gleichung erhält man dann durch Anwendung des Umkehroperators zum Operator (6). Das Ergebnis der Anwendung eines solchen Operators auf einfache Funktionen wie trigonometrische Funktionen, Exponentialfunktionen, Treppenfunktionen und deren Linearkombinationen kann man durch unmittelbares Ausrechnen finden. Nach dieser Methode kann man daher die Lösung einer linearen Gleichung mit konstanten Koeffizienten automatisch aufschreiben, wenn die rechte Seite eine Linearkombination derartiger Funktionen ist. Es ist klar, daß die Operatorenmethode in gewisser Weise eine Anwendung der Laplaceformation darstellt (die ja eine bestimmte Zuordnung zwischen der Algebra der Differentialoperatoren (6) und der Algebra der Polynome herstellt). Man kann die Laplacetransformation als Rechtfertigung dieser Methode ansehen, die in der technischen Literatur oft in Form eines „Rezeptes“ erscheint.

8.7. Die Fourier-Stieltjes-Transformation

8.7.1. Definition der Fourier-Stieltjes-Transformation. Wir kehren noch einmal zur Fouriertransformation im Raum $L_1(-\infty, \infty)$ zurück:

$$g(\lambda) = \int_{-\infty}^{\infty} e^{-i\lambda x} f(x)\, dx.$$

Diese Formel kann man auch als Riemann-Stieltjes-Integral

$$g(\lambda) = \int_{-\infty}^{\infty} e^{-i\lambda x}\, dF(x) \tag{1}$$

schreiben, wobei

$$F(x) = \int_{-\infty}^{x} f(t)\, dt \tag{2}$$

eine absolut stetige Funktion beschränkter Variation auf der ganzen reellen Achse ist (die Totalvariation ist $\int_{-\infty}^{\infty} |f(x)|\, dx$). Formel (1) besitzt jedoch nicht nur für Funktionen der Form (2) einen Sinn, sondern für beliebige Funktionen beschränkter Variation. Das Integral

$$g(\lambda) = \int_{-\infty}^{\infty} e^{-i\lambda x}\, dF(x),$$

in dem F eine beliebige Funktion von beschränkter Variation auf der reellen Achse ist, nennen wir *Fourier-Stieltjes-Transformierte* der Funktion F.

Für die Fourier-Stieltjes-Transformation bleiben eine Reihe von Eigenschaften erhalten, die wir früher für die gewöhnliche Fouriertransformation nachgewiesen

haben. So gilt zum Beispiel folgendes: Die durch das Integral (1) definierte Funktion g ist auf der ganzen Geraden stetig und beschränkt.

Die Beschränktheit ist klar. Um die Stetigkeit zu zeigen, schätzen wir die Differenz zweier Funktionswerte ab:

$$|g(\lambda_1) - g(\lambda_2)| \leqq \int_{-N}^{N} |e^{-i\lambda_1 x} - e^{-i\lambda_2 x}| \, dF(x)$$
$$+ \int_{|x|>N} |e^{-i\lambda_1 x} - e^{-i\lambda_2 x}| \, dF(x).$$

Den zweiten Summanden kann man hier (unabhängig von λ_1 und λ_2) beliebig klein machen, wenn man N hinreichend groß wählt. Der erste Summand strebt bei festem N für $\lambda_1 - \lambda_2 \to 0$ gegen Null.

Es lassen sich jedoch nicht alle Eigenschaften der Fouriertransformation auf die Fourier-Stieltjes-Transformation übertragen. So strebt die Fourier-Stieltjes-Transformierte im allgemeinen für $|\lambda| \to \infty$ nicht gegen Null. Ist zum Beispiel

$$F(x) = \begin{cases} 0 & \text{für} \quad x \leqq 0, \\ 1 & \text{für} \quad x > 0, \end{cases}$$

so gilt

$$g(\lambda) = \int_{-\infty}^{\infty} e^{i\lambda x} \, dF(x) = 1.$$

Analog ist die Fourier-Stieltjes-Transformierte der Funktion, die für $x \leqq x_0$ den Wert 0 und für $x > x_0$ den Wert 1 annimmt, die periodische Funktion $e^{ix_0\lambda}$ von λ.

Ist F eine Treppenfunktion mit den Sprungstellen

$$n = 0, \pm 1, \pm 2, \ldots$$

und den zugehörigen Sprunghöhen

$$\ldots, a_{-1}, a_0, a_1, \ldots, a_n, \ldots \qquad \left(\sum_n |a_n| < \infty\right),$$

so ist

$$\int_{-\infty}^{\infty} e^{-i\lambda x} \, dF(x) = \sum_n a_n e^{-in\lambda}$$

eine periodische Funktion mit der Periode 2π. Besitzt die Funktion F Sprünge der Höhe a_n in den Punkten x_n einer beliebigen Folge von (im allgemeinen inkommensurablen) Zahlen, so nimmt die Fourier-Stieltjes-Transformierte von F die Gestalt

$$\sum_n a_n e^{-ix_n\lambda}$$

an. Funktionen dieses Typs gehören zu den sogenannten *fastperiodischen* Funktionen.

8.7.2. Anwendung der Fourier-Stieltjes-Transformation in der Wahrscheinlichkeitsrechnung. Für integrierbare Funktionen auf dem Intervall $(-\infty, \infty)$ haben wir in 8.4. den Begriff der Faltung eingeführt:

$$f(x) = f_1 * f_2(x) = \int_{-\infty}^{\infty} f_1(x-\xi)\, f_2(\xi)\, d\xi. \tag{3}$$

Setzen wir

$$F(x) = \int_{-\infty}^{x} f(t)\, dt, \qquad F_1(x) = \int_{-\infty}^{x} f_1(t)\, dt \quad \text{und} \quad F_2(x) = \int_{-\infty}^{x} f_2(t)\, dt,$$

so erhalten wir nach Integration von (3) die Gleichung

$$F(x) = \int_{-\infty}^{x} f(t)\, dt = \int_{-\infty}^{x} dt \int_{-\infty}^{\infty} f_1(t-\xi)\, f_2(\xi)\, d\xi$$

$$= \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{x} f_1(t-\xi)\, dt \right\} f_2(\xi)\, d\xi = \int_{-\infty}^{\infty} F_1(x-\xi)\, dF_2(\xi)$$

(die Vertauschung der Integrationsreihenfolge ist hier wegen der absoluten Integrierbarkeit von f nach dem Satz von Fubini möglich). Die oben erhaltene Beziehung

$$F(x) = \int_{-\infty}^{\infty} F_1(x-\xi)\, dF_2(\xi)$$

ordnet den Funktionen F_1 und F_2 die Funktion F zu. Das Lebesgue-Stieltjes-Integral auf der rechten Seite existiert jedoch nicht nur für absolut stetige Funktionen, sondern sogar für je zwei Funktionen von beschränkter Variation. Der Ausdruck

$$F(x) = \int_{-\infty}^{\infty} F_1(x-\xi)\, dF_2(\xi) \tag{4}$$

mit zwei beliebigen Funktionen F_1 und F_2 von beschränkter Variation wird *Faltung dieser beiden Funktionen* genannt und mit $F_1 * F_2$ bezeichnet.

Wir wollen zeigen, daß der Ausdruck (4) eine auf der ganzen reellen Achse definierte Funktion von beschränkter Variation darstellt.[1]) Ist F_1 eine Funktion von beschränkter Variation, so ist F_1 auch Borel-meßbar, und folglich existiert das Integral (4) für alle x. Ferner ist

$$|F(x_1) - F(x_2)| = \left| \int_{-\infty}^{\infty} \big(F_1(x_1-\xi) - F_1(x_2-\xi)\big)\, dF_2(\xi) \right|$$

$$\leqq \int_{-\infty}^{\infty} |F_1(x_1-\xi) - F_1(x_2-\xi)|\, d\big(\operatorname{var} F_2(\xi)\big),$$

[1]) Im Buch von V. I. Glivenko (Das Stieltjesintegral [russ.], Gostechisdat, Moskau 1936) ist eine elementare Konstruktion angegeben, mit deren Hilfe man (4) ohne Benutzung des Maßbegriffs einen Sinn zuordnen kann.

woraus

$$V[F] \leqq V[F_1] \cdot V[F_2]$$

folgt. Also ist F eine Funktion von beschränkter Variation.

Satz 1. *Ist F die Faltung der Funktionen F_1 und F_2 von beschränkter Variation und sind g, g_1 und g_2 die entsprechenden Fourier-Stieltjes-Transformierten, so gilt*

$$g(\lambda) = g_1(\lambda) \cdot g_2(\lambda).$$

Beweis. Es sei $F = F_1 * F_2$ und

$$a_0 = x_0, x_1, \ldots, x_n = b$$

eine Zerlegung des Intervalls $[a, b]$. Dann ist für jedes λ

$$\int_a^b e^{-i\lambda x} dF(x) = \lim_{\max \Delta x_k \to 0} \sum_{k=1}^n e^{-i\lambda x_k}\big(F(x_k) - F(x_{k-1})\big)$$

$$= \lim_{\max \Delta x_k \to 0} \int_{-\infty}^{\infty} \sum_{k=1}^n e^{-i\lambda(x_k-\xi)}\big(F_1(x_k - \xi) - F_1(x_{k-1} - \xi)\big)\, e^{-i\lambda\xi}\, dF_2(\xi),$$

d. h.

$$\int_a^b e^{-i\lambda x}\, dF(x) = \int_{-\infty}^{\infty} \left\{ \int_{a-\xi}^{b-\xi} e^{-i\lambda x}\, dF_1(x) \right\} e^{-i\lambda\xi}\, dF_2(\xi).$$

Durch Grenzübergang $a \to \infty$ und $b \to \infty$ erhält man hieraus

$$\int_{-\infty}^{\infty} e^{-i\lambda x}\, dF(x) = \int_{-\infty}^{\infty} e^{-i\lambda x}\, dF_1(x) \cdot \int_{-\infty}^{\infty} e^{-i\lambda\xi}\, dF_2(\xi),$$

d. h.

$$g(\lambda) = g_1(\lambda) \cdot g_2(\lambda).$$

Die Tatsache, daß die Fourier-Stieltjes-Transformierte die Faltung von Funktionen in eine Multiplikation verwandelt, wird oft in der Wahrscheinlichkeitsrechnung verwendet (*Methode der charakteristischen Funktionen*). Dort besteht oft die Notwendigkeit, die Summe zufälliger unabhängiger Summanden zu betrachten. Sind ξ und η zwei unabhängige Zufallsgrößen und F_1 und F_2 die zugehörigen Verteilungsfunktionen, so entspricht der Zufallsgröße $\xi + \eta$ die Verteilungsfunktion

$$F = F_1 * F_2.$$

Der Übergang von Verteilungsfunktionen zu ihren Fourier-Stieltjes-Transformierten, den sogenannten charakteristischen Funktionen, gestattet also die Ersetzung der Faltung durch die einfachere und bequemere Multiplikation.

Aufgaben

1. Es ist zu beweisen, daß die Fourier-Stieltjes-Transformation folgende Eigenschaft besitzt: Wenn die Funktion F linksseitig stetig ist und ihre Fourier-Stieltjes-Transformierte identisch verschwindet, ist $F(x) = \text{const}$.

2. Man beweise, daß die Faltung von Funktionen beschränkter Variation kommutativ und assoziativ ist.

8.8. Die Fouriertransformation für Distributionen

Die Fouriertransformation im üblichen Sinne ist nur für Funktionen erklärt, die auf der ganzen reellen Achse integrierbar sind. Wie bereits erwähnt, wird die Anwendung der Fouriertransformation auf Differentialgleichungen und andere Fragen dadurch stark eingeschränkt. Den Anwendungsbereich der Fouriertransformation kann man jedoch wesentlich erweitern, indem man den Begriff der Fouriertransformation von Distributionen einführt. Wir wollen die Grundgedanken dieser Konstruktion darlegen.

Zunächst erinnern wir an den Raum S_∞ der Funktionen, die auf der ganzen reellen Achse beliebig oft differenzierbar sind und im Unendlichen zusammen mit ihren Ableitungen schneller als jede Potenz von $1/|x|$ fallen (vgl. 4.4.). Wir nehmen S_∞ als *Raum der Grundfunktionen* und betrachten den zugehörigen *Raum von Distributionen* $S_\infty{}^*$.

Im Raum $S_\infty{}^*$ definieren wir jetzt die Fouriertransformation. Dazu erinnern wir zunächst daran, daß die Fouriertransformation (im gewöhnlichen Sinne) den Raum S_∞ in sich überführt: Ist $\varphi \in S_\infty$, so ist auch $F[\varphi] \in S_\infty$; dabei ist F eine eineindeutige Abbildung von S_∞ auf sich. Davon ausgehend gelangt man zu folgender Definition: Das lineare Funktional $g \in S_\infty{}^*$, das durch die Formel

$$(g, \psi) = 2\pi(f, \varphi) \quad \text{mit} \quad \psi = F[\varphi] \tag{1}$$

definiert wird, nennen wir die *Fouriertransformierte der Distribution* $f \in S_\infty{}^*$. Diese Formel kann man auch in der Gestalt

$$(Ff, \psi) = 2\pi(f, \varphi) = 2\pi(f, F^{-1}\psi)$$

schreiben. Die Fouriertransformierte des Funktionals $f \in S_\infty{}^*$ ist also das Funktional, das jedem Element $\psi \in S_\infty$ den (mit 2π multiplizierten) Wert von f auf dem Element $\varphi = F^{-1}\psi$ zuordnet. Dabei ist F^{-1} die inverse Fouriertransformation.

Gleichung (1) definiert ein Funktional auf ganz S_∞. Denn wenn φ den Raum S_∞ durchläuft, so durchläuft auch $\psi = F[\varphi]$ ganz S_∞. Die Linearität und die Stetigkeit dieses Funktionals lassen sich unmittelbar überprüfen.

Zu den Elementen aus $S_\infty{}^*$ gehören auch alle integrierbaren Funktionen. Für diese Funktionen stimmt die eben angegebene Definition der Fouriertransformation mit der üblichen Definition überein, denn nach dem Satz von Plancherel gilt für $f \in S_\infty$,

$\varphi \in S_\infty$ und $g = F[f]$, $\psi = F[\varphi]$ die Gleichung

$$2\pi(f, \varphi) = (g, \psi). \tag{2}$$

Dabei existiert zu gegebenem f bis auf Äquivalenz nur eine Funktion g, die dieser Gleichung für alle $\varphi \in S_\infty$ genügt. Mit Hilfe eines entsprechenden Grenzüberganges kann man unschwer zeigen, daß Gleichung (2) auch für beliebige $f \in L_1(-\infty, \infty)$ gilt. Somit stellt die Fouriertransformation von Distributionen eine Fortsetzung der klassischen Fouriertransformation auf eine größere Klasse von Objekten dar.

Beispiele

1. Es sei $f(x) = c = \text{const}$. Dann ist

$$2\pi(f, \varphi) = 2\pi \int_{-\infty}^{\infty} c\varphi(x)\, dx = 2\pi c\psi(0) \qquad (\psi = F[\varphi]),$$

d. h., die Fouriertransformierte einer Konstanten ist die δ-Funktion, multipliziert mit dieser Konstanten und 2π.

2. Es sei $f(x) = e^{iax}$. Dann ist

$$2\pi(f, \varphi) = 2\pi \int_{-\infty}^{\infty} e^{-iax}\varphi(x)\, dx = 2\pi\psi(-a),$$

d. h., die Fouriertransformierte der Funktion e^{iax} ist die verschobene δ-Funktion $\delta(x + a)$, multipliziert mit 2π.

3. Es sei $f(x) = x^2$, dann erhalten wir aus der Gleichung

$$\psi''(\lambda) = -\int_{-\infty}^{\infty} x^2\varphi(x)\, e^{-i\lambda x}\, dx$$

nach Multiplikation mit 2π für $x = 0$

$$2\pi\big(x^2, \varphi(x)\big) = -2\pi\psi''(0),$$

d. h., die Fouriertransformation der Funktion x^2 ist die zweite Ableitung der δ-Funktion, multipliziert mit -2π.

Wir wollen noch einige abschließende Bemerkungen machen.

Wir haben die Fouriertransformation für Distributionen über S_∞ erklärt. Man könnte aber auch jeden anderen Grundraum nehmen, z. B. den Raum K aller unendlich oft differenzierbaren finiten Funktionen. Für jede Funktion $\varphi \in K$ existiert die Fouriertransformierte (im gewöhnlichen Sinne) und stellt, wie man leicht nachprüfen kann, eine ganze analytische, exponentiell wachsende Funktion dar. Das bedeutet, daß die Fouriertransformation ein eineindeutiger linearer Operator ist, der den Raum K auf den folgenden Raum Z abbildet. Die Elemente von Z sind alle ganzen

analytischen Funktionen ψ, die für $q = 1, 2, \ldots$ Ungleichungen der Form

$$|s|^q\,|\psi(s)| \leqq C_q e^{a|\tau|}$$

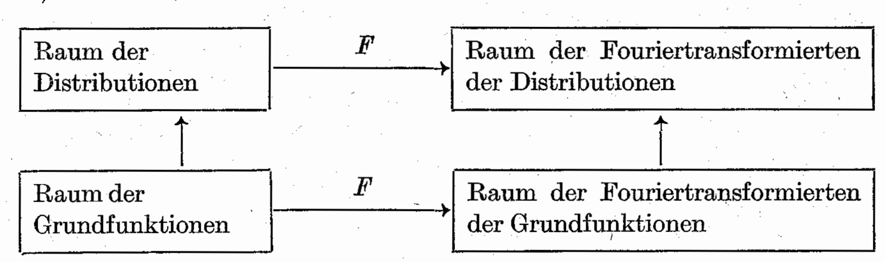
mit $\tau = \operatorname{Im} s$ und zwei von ψ abhängigen Konstanten C_q und a genügen. Im Raum K hatten wir einen Konvergenzbegriff erklärt. Durch die Transformation F, die K auf Z abbildet, wird dann in Z ein Konvergenzbegriff induziert. Eine Folge $\{\psi_n\}$ konvergiert in Z gegen ψ, wenn die Beziehung $\varphi_n \to \varphi$ für die entsprechenden Urbilder gilt. Man kann diesen Konvergenzbegriff übrigens auch ohne Verwendung des Raumes K formulieren.[1])

Jetzt sei f ein beliebiges Element aus K^*. Setzen wir

$$(g, \psi) = 2\pi(f, \varphi) \quad \text{mit} \quad \psi = F[\varphi],$$

so wird der Distribution f das lineare Funktional g zugeordnet. Dieses Funktional g nennen wir die *Fouriertransformierte* von f. Somit ist die Fouriertransformierte einer Distribution f über dem Grundraum K eine Distribution über Z, d. h. über dem Raum, in den K durch die Fouriertransformation (im gewöhnlichen Sinne) übergeführt wird.

Die gleiche Konstruktion ist auch für Distributionen über jedem anderen Raum von Grundfunktionen möglich. Dabei ergibt sich jedesmal ein Schema, das vier Räume umfaßt: einen Ausgangsraum von Grundfunktionen, die Gesamtheit der Fouriertransformierten dieser Funktionen (d. h. einen zweiten Raum von Grundfunktionen) und zwei duale Räume.



Dieses Schema enthält nur zwei Räume, wenn man S_∞ als Grundraum nimmt, da dieser Raum durch die Fouriertransformation in sich übergeführt wird.

Der Begriff der Fouriertransformation für Distributionen findet in der Theorie der partiellen Differentialgleichungen breite Anwendung. Der Leser kann sich mit diesen Fragen zum Beispiel im Buch von G. E. Šilov [60] vertraut machen.[2])

---

[1]) Und zwar gilt $\psi_n \to 0$ in Z, wenn für feste C_q $(q = 1, 2, \ldots)$ und a die Ungleichungen

$$|s^q \psi_n(s)| \leqq C_q e^{a|\tau|}$$

erfüllt sind und ψ_n auf jedem endlichen Intervall der reellen Achse gleichmäßig gegen Null strebt.

[2]) Anm. d. Übers.: Man vergleiche auch [56], [65], [69].

9. Lineare Integralgleichungen

9.1. Grundlegende Definitionen. Einige Probleme, die auf lineare Integralgleichungen führen

9.1.1. Integralgleichungstypen. Eine Gleichung, in der die unbekannte Funktion auch unter dem Integralzeichen steht, heißt *Integralgleichung*. So ist z. B. die Gleichung

$$\varphi(s) = \int_a^b K(s, t)\, \varphi(t)\, dt + f(s), \tag{1}$$

in der f und K bekannte Funktionen und φ eine gesuchte Funktion ist, eine Integralgleichung. Die Variablen s und t durchlaufen hier ein festes Intervall $[a, b]$. Eine charakteristische Besonderheit der Gleichung (1) ist ihre Linearität: Die unbekannte Funktion geht in die Gleichung linear ein. Eine Reihe von Problemen führt auch auf nichtlineare Integralgleichungen, z. B. auf Gleichungen von der Form

$$\varphi(s) = \int_a^b K(s, t)\, g\big(\varphi(t), t\big)\, dt,$$

wobei K und g gegebene Funktionen sind. Wir werden uns jedoch im weiteren auf die Betrachtung linearer Gleichungen beschränken.

Einzelne Integralgleichungen wurden schon zu Anfang des vorigen Jahrhunderts betrachtet. So untersuchte Abel bereits 1823 die Gleichung

$$f(s) = \int_0^s \frac{\varphi(t)}{(s - t)^\alpha}\, dt \qquad (0 < \alpha < 1,\, f(0) = 0),$$

die noch heute seinen Namen trägt. Dabei ist f eine gegebene und φ eine gesuchte Funktion. Abel zeigte, daß die Lösung dieser Gleichung die Form

$$\varphi(t) = \frac{\sin \pi\alpha}{\pi} \int_0^t \frac{f'(s)}{(t - s)^{1-\alpha}}\, ds$$

hat. Eine allgemeine Theorie der linearen Integralgleichungen wurde aber erst am Ende des 19. und zu Beginn des 20. Jahrhunderts auf der Grundlage der Arbeiten von Volterra, Fredholm und Hilbert geschaffen.

Eine Gleichung der Form (1) heißt *Fredholmsche Integralgleichung zweiter Art* (vgl. 2.4.4.); eine Gleichung der Form

$$\int_a^b K(s, t)\, \varphi(t)\, dt + f(s) = 0 \tag{2}$$

(in welcher die unbekannte Funktion φ nur unter dem Integralzeichen vorkommt) heißt *Fredholmsche Integralgleichung erster Art.*

Die oben erwähnte Abelsche Integralgleichung zählt zu den sogenannten Volterraschen Integralgleichungen. Die allgemeine Form dieser Integralgleichungen ist

$$\int_a^s K(s, t)\, \varphi(t)\, dt = f(s) \tag{3}$$

(*Volterrasche Integralgleichung erster Art*) oder

$$\varphi(s) = \int_a^s K(s, t)\, \varphi(t)\, dt + f(s) \tag{4}$$

(*Volterrasche Integralgleichung zweiter Art*). Offensichtlich kann man die Volterraschen Integralgleichungen als spezielle Fredholmsche Integralgleichungen auffassen, in denen die Funktion K der Bedingung

$$K(s, t) = 0 \qquad \text{für } t > s$$

genügt. Es ist jedoch zweckmäßig, die Volterraschen Integralgleichungen als spezielle Klasse abzutrennen, da sie eine Reihe von wesentlichen Eigenschaften besitzen, die bei beliebigen Fredholmschen Integralgleichungen fehlen.

Wenn in den Gleichungen (1), (2) oder (3) die Funktion f gleich 0 ist, nennt man diese Gleichungen *homogen*; anderenfalls heißen sie *inhomogen.*

9.1.2. Beispiele für Probleme, die auf Integralgleichungen führen. In den nächsten Abschnitten dieses Kapitels werden wir grundlegende Eigenschaften linearer Integralgleichungen betrachten. Zuvor beschreiben wir aber einige Aufgaben, die auf solche Gleichungen führen.

1. *Die Gleichgewichtslage einer belasteten Saite.* Wir betrachten eine Saite, d. h. einen elastischen materiellen Faden der Länge l, der sich frei verbiegen kann, aber einer Dehnung einen Widerstand proportional zur Größe dieser Dehnung entgegensetzt. Die Enden der Saite seien in den Punkten $x = 0$ und $x = l$ fest eingespannt. Dann stimmt die Gleichgewichtslage der Saite mit dem Intervall $0 \leqq x \leqq l$ der x-Achse überein. Nehmen wir jetzt an, daß im Punkt $x = \xi$ auf die Saite eine vertikalgerichtete Kraft $P = P_\xi$ einwirkt, so wird die Saite unter dem Einfluß der Kraft von der Gleichgewichtslage abweichen und eine Form annehmen, wie sie in Abb. 23 dargestellt ist.

Ist δ die Größe der Auslenkung der Saite im Punkt ξ unter dem Einfluß der Kraft P_ξ und die Kraft P_ξ klein im Vergleich zur Spannung T_0 der unbelasteten Saite, dann kann man die Spannung der belasteten Saite auch mit T_0 annehmen und aus

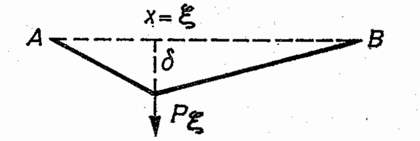


Abb. 23

der Gleichgewichtsbedingung folgern, daß

$$T_0 \frac{\delta}{\xi} + T_0 \frac{\delta}{l - \xi} = P_\xi,$$

d. h.

$$\delta = \frac{(l - \xi)\, \xi}{T_0 l} P_\xi$$

ist. Bezeichnen wir $u(x)$ die Auslenkung der Saite im Punkt x unter dem Einfluß der Kraft P_ξ, dann ist

$$u(x) = P_\xi G(x, \xi)$$

mit

$$G(x, \xi) = \begin{cases} \dfrac{x(l - \xi)}{T_0 l} & \text{für } 0 \leqq x \leqq \xi, \\ \dfrac{(l - x)\, \xi}{T_0 l} & \text{für } \xi \leqq x \leqq l. \end{cases}$$

Aus dieser Formel folgt sofort $G(x, \xi) = G(\xi, x)$. Nehmen wir nun an, daß längs der ganzen Saite eine stetig verteilte Kraft mit der Dichte $p(\xi)$ wirkt und diese Kraft klein im Vergleich zur Spannung der Saite ist (d. h. die Deformation linear von der Kraft abhängt), so wird die Auslenkung $u(x)$ der Saite im Punkt x durch die Formel

$$u(x) = \int_0^l G(x, \xi)\, p(\xi)\, d\xi \tag{5}$$

beschrieben. Damit kann bei vorgegebener Belastung mit Hilfe von (5) die Form der Saite berechnet werden, die diese unter dem Einfluß der Kraft annimmt.

Die umgekehrte Aufgabe, d. h. die Bestimmung einer Kraftverteilung p, unter deren Einfluß die Saite eine vorgegebene Form $u(x)$ annimmt, führt gemäß (5) zu einer Fredholmschen Integralgleichung erster Art.

2. *Freie und erzwungene Schwingungen einer Saite.* Wir betrachten jetzt eine schwingende Saite. Dabei sei $u(x, t)$ die Auslenkung des Saitenpunktes mit der Ab-

szisse x zum Zeitpunkt t, und ϱ sei die (lineare) Dichte der Saite.[1]) Dann wirkt auf ein Saitenstück der Länge dx eine Trägheitskraft der Größe

$$-\frac{\partial^2 u(x, t)}{\partial t^2} \varrho\, dx,$$

woraus

$$p(\xi) = -\frac{\partial^2 u(\xi, t)}{\partial t^2} \varrho$$

folgt. Setzen wir diesen Ausdruck an Stelle von $p(\xi)$ in die Formel (5) ein, so erhalten wir

$$u(x, t) = -\int_0^l G(x, \xi)\, \varrho \frac{\partial^2 u(\xi, t)}{\partial t^2}\, d\xi. \tag{6}$$

Nehmen wir weiter an, daß die Saite harmonische Schwingungen mit der festen Kreisfrequenz ω und der nur von x abhängigen Amplitude $u(x)$ ausführt, d. h., ist

$$u(x, t) = u(x) \sin \omega t,$$

dann erhalten wir entsprechend (6) für u die Integralgleichung

$$u(x) = \varrho\omega^2 \int_0^l G(x, \xi)\, u(\xi)\, d\xi. \tag{7}$$

Wenn die Saite keine freien Schwingungen, sondern unter dem Einfluß einer äußeren Kraft erzwungene Schwingungen ausführt, dann zeigt eine ähnliche Rechnung, daß die Gleichung der harmonischen Schwingungen die Form

$$u(x) = \varrho\omega^2 \int_0^l G(x, \xi)\, u(\xi)\, d\xi + f(x)$$

hat, d. h. die Amplitude $u(x)$ einer inhomogenen Fredholmschen Integralgleichung zweiter Art genügt.

3. *Zurückführung von Differentialgleichungen auf Integralgleichungen.* In manchen Fällen ist es zweckmäßig, die Lösung einer Differentialgleichung auf die Lösung einer Integralgleichung zurückzuführen. Wenn z. B. die Existenz und Eindeutigkeit einer Lösung der Differentialgleichung

$$y' = f(x, y)$$

mit den Anfangsbedingungen $y(x_0) = y_0$ bewiesen werden soll, dann ist es, wie wir in Kapitel 2 sahen, nützlich, zur entsprechenden (nichtlinearen) Integralgleichung

$$y = y_0 + \int_{x_0}^{x} f(\xi, y)\, d\xi$$

---

[1]) Wir nehmen an, daß $\varrho = \text{const}$ ist, obwohl dies für die weitere Rechnung nicht wesentlich ist.

überzugehen. Eine solche Überführung ist auch für Differentialgleichungen höherer als erster Ordnung möglich. Wir betrachten als Beispiel die Differentialgleichung zweiter Ordnung

$$y'' + f(x)\,y = 0.$$

Setzen wir $f(x) = \varrho^2 - \sigma(x)$, wobei $\varrho = \text{const}$ ist, so erhält diese Gleichung die Form

$$y'' + \varrho^2 y = \sigma(x)\,y. \tag{8}$$

Da sich bekanntlich die Lösung der Gleichung

$$y'' + \varrho^2 y = g(x)$$

in der Form

$$y(x) = \cos \varrho(x - a) + \frac{1}{\varrho} \int_a^x \sin \varrho(x - \xi) \cdot g(\xi)\,d\xi$$

darstellen läßt, kann die Lösung der Gleichung (8) auf die Lösung der Integralgleichung

$$y(x) - \frac{1}{\varrho} \int_a^x \sigma(\xi) \sin \varrho(x - \xi)\,y(\xi)\,d\xi = \cos \varrho(x - a)$$

zurückgeführt werden.

9.2. Fredholmsche Integralgleichungen

9.2.1. Der Fredholmsche Integraloperator. In diesem Abschnitt betrachten wir Fredholmsche Integralgleichungen zweiter Art, d. h. Gleichungen der Form

$$\varphi(s) = \int_a^b K(s, t)\,\varphi(t)\,dt + f(s). \tag{1}$$

Alle hier und im weiteren auftretenden Funktionen nehmen wir als komplexwertig an. Bezüglich der Funktion K, dem sogenannten *Kern* der Integralgleichung, setzen wir voraus, daß sie meßbar ist und zur Klasse L_2 über dem Quadrat $a \leqq s, t \leqq b$ gehört, d. h., daß

$$\int_a^b \int_a^b |K(s, t)|^2\,ds\,dt < \infty \tag{2}$$

ist. Solche Kerne werden auch als *Hilbert-Schmidtsche Kerne* bezeichnet. Das freie Glied f der Gleichung (1) sei eine gegebene, φ eine gesuchte Funktion aus $L_2[a, b]$.

Definieren wir, ausgehend von Gleichung (1), durch

$$A\varphi = \int_a^b K(s, t)\, \varphi(t)\, dt = \psi(s) \tag{3}$$

einen linearen Operator A, so kann die Untersuchung von (1) auf die Untersuchungen der Eigenschaften dieses Operators zurückgeführt werden. Ein solcher Operator (3) wird allgemein als *Fredholmscher Operator* und, wenn der Kern die Bedingung (2) erfüllt, auch als *Hilbert-Schmidt-Operator* bezeichnet.

Satz 1. *Erfüllt $K(s, t)$ die Bedingung (2), dann definiert (3) im Raum $L_2[a, b]$ einen kompakten linearen Operator A, dessen Norm der Abschätzung*

$$\|A\| \leqq \sqrt{\int_a^b \int_a^b |K(s, t)|^2\, ds\, dt} \tag{4}$$

genügt.

Beweis. Wir bemerken zuerst, daß auf Grund der vorausgesetzten Eigenschaft (2) das Integral

$$\int_a^b |K(s, t)|^2\, dt$$

nach dem Satz von Fubini für fast alle s existiert. $K(s, t)$ gehört also als Funktion von t für fast alle s zu $L_2[a, b]$. Da das Produkt quadratisch integrierbarer Funktionen integrierbar ist, existiert das Integral in (3) für fast alle s, d. h., die Funktion $\psi(s)$ ist fast überall definiert. Wir zeigen nun, daß $\psi \in L_2[a, b]$ ist. Nach der Schwarzschen Ungleichung gilt für fast alle s

$$|\psi(s)|^2 = \left|\int_a^b K(s, t)\, \varphi(t)\, dt\right|^2 \leqq \int_a^b |K(s, t)|^2\, dt \cdot \int_a^b |\varphi(t)|^2\, dt$$

$$\leqq \|\varphi\|^2 \cdot \int_a^b |K(s, t)|^2\, dt.$$

Integrieren wir diese Ungleichung über s und verwandeln das iterierte Integral von $|K(s, t)|^2$ in ein zweifaches, so erhalten wir die Ungleichung

$$\|A\varphi\|^2 = \int_a^b |\psi(s)|^2\, ds \leqq \|\varphi\|^2 \int_a^b \int_a^b |K(s, t)|^2\, ds\, dt,$$

aus welcher die Integrierbarkeit von $|\psi(s)|^2$ und die Abschätzung (4) für die Norm des Operators A folgen. Es bleibt nun nur noch die Kompaktheit des Operators A nachzuweisen. Dazu betrachten wir ein vollständiges Orthogonalsystem $\{\psi_n\}$ in $L_2[a, b]$. Bilden wir alle möglichen Produkte $\psi_m(s)\, \psi_n(t)$, so ergibt sich nach Satz 1 aus 7.3. ein

vollständiges System im Raum $L_2([a, b] \times [a, b])$ und damit eine Entwicklung für den Kern $K(s, t)$:

$$K(s, t) = \sum_{m,n=1}^{\infty} a_{mn}\psi_m(s)\,\psi_n(t).$$

Setzen wir

$$K_N(s, t) = \sum_{m,n=1}^{N} a_{mn}\psi_m(s)\,\psi_n(t)$$

und bezeichnen mit A_N den Operator, der dem Kern $K_N(s, t)$ entspricht, dann ist dieser Operator kompakt, weil er ganz $L_2[a, b]$ in einen endlichdimensionalen Unterraum abbildet (in Kapitel 4 nannten wir solche Operatoren endlichdimensionale Operatoren). Für $\varphi \in L_2[a, b]$ gilt nämlich

$$A_N\varphi = \int_a^b K_N(s, t)\,\varphi(t)\,dt = \sum_{m,n=1}^{N} a_{mn}\psi_m(s) \int_a^b \varphi(t)\,\psi_n(t)\,dt$$

$$= \sum_{m=1}^{N} \psi_m(s) \sum_{n=1}^{N} a_{mn}b_n,$$

wobei

$$b_n = \int_a^b \varphi(t)\,\psi_n(t)\,dt$$

ist, d. h., der Operator A_N führt jedes Element $\varphi \in L_2[a, b]$ in ein Element eines endlichdimensionalen Raumes über, der durch die Vektoren $\psi_1, \ldots, \psi_N$ erzeugt wird. Da $K_N(s, t)$ andererseits die N-te Partialsumme der Fourierreihe für die Funktion $K(s, t)$ ist, gilt

$$\int_a^b\int_a^b |K(s, t) - K_N(s, t)|^2\,ds\,dt \to 0 \quad \text{für } N \to \infty.$$

Hieraus folgt, wenn wir die Abschätzung (4) auf den Operator $A - A_N$ anwenden,

$$\|A - A_N\| \to 0 \qquad \text{für } N \to \infty.$$

Benutzen wir nun Satz 1 aus 4.6., nach dem der Grenzwert einer konvergenten Folge kompakter Operatoren ebenfalls kompakt ist, so ergibt sich die Kompaktheit des Operators A. Damit ist der Satz bewiesen.

Bemerkungen

1. Beim Beweis des Satzes 1 haben wir gezeigt, daß jeder Hilbert-Schmidt-Operator als Grenzwert (im Sinne der Normkonvergenz) einer Folge endlichdimensionaler Integraloperatoren dargestellt werden kann.

2. A_1 und A_2 seien zwei Operatoren der Form (3) und $K_1(s, t)$, $K_2(s, t)$ die ihnen entsprechenden Kerne. Wenn die Operatoren A_1 und A_2 gleich sind, d. h. $A_1\varphi = A_2\varphi$

für alle $\varphi \in L_2[a, b]$ gilt, dann ist fast überall auf $[a, b] \times [a, b]$

$$K_1(s, t) = K_2(s, t),$$

denn aus

$$A_1\varphi - A_2\varphi = \int_a^b \big(K_1(s, t) - K_2(s, t)\big)\, \varphi(t)\, dt = 0$$

für alle $\varphi \in L_2[a, b]$ ergibt sich

$$\int_a^b |K_1(s, t) - K_2(s, t)|^2\, dt = 0$$

für fast alle $s \in [a, b]$, und damit ist auch

$$\int_a^b \int_a^b |K_1(s, t) - K_2(s, t)|^2\, ds\, dt = 0,$$

woraus unmittelbar die Behauptung folgt. Identifizieren wir wie üblich zueinander äquivalente integrierbare Funktionen, so können wir nun feststellen, daß zwischen den Integraloperatoren und den Kernen eine umkehrbar eindeutige Zuordnung besteht.

Satz 2. *Ist A ein Hilbert-Schmidt-Operator, der durch den Kern $K(s, t)$ definiert wird, dann wird der zu A adjungierte Operator A^* durch den „adjungierten" Kern $\overline{K(t, s)}$ definiert.*

Beweis. Unter Benutzung des Satzes von Fubini erhält man

$$\begin{aligned}(Af, g) &= \int_a^b \left\{ \int_a^b K(s, t)\, f(t)\, dt \right\} \overline{g(s)}\, ds \\ &= \int_a^b \int_a^b K(s, t)\, f(t)\, \overline{g(s)}\, dt\, ds \\ &= \int_a^b \left\{ \int_a^b K(s, t)\, \overline{g(s)}\, ds \right\} f(t)\, dt \\ &= \int_a^b f(t)\, \overline{\left\{ \int_a^b \overline{K(s, t)}\, g(s)\, ds \right\}}\, dt,\end{aligned}$$

woraus die Behauptung des Satzes folgt.

Insbesondere ist also ein Operator A der Form (3) in $L_2[a, b]$ genau dann selbstadjungiert, d. h., es gilt $A^* = A$, wenn $\overline{K(s, t)} = K(t, s)$ ist. Wird nur der reelle Raum $L_2[a, b]$ (damit auch nur reelle Kerne) betrachtet, dann lautet die Bedingung

für die Selbstadjungiertheit des Operators A

$$K(s, t) = K(t, s).$$

Bemerkung. Wir haben in diesem Abschnitt nur Integraloperatoren im Raum $L_2[a, b]$ betrachtet. Alle bisherigen und auch alle weiteren Resultate sind ohne Änderungen auf den allgemeineren Fall übertragbar, bei dem an Stelle des Intervalls $[a, b]$ eine beliebige andere Menge mit Maß betrachtet wird.

9.2.2. Integralgleichungen mit symmetrischem Kern. Wir betrachten eine Fredholmsche Integralgleichung zweiter Art

$$\varphi(s) = \int_a^b K(s, t)\, \varphi(t)\, dt + f(s), \tag{5}$$

deren Kern den Bedingungen

1. $\int_a^b \int_a^b |K(s, t)|^2\, ds\, dt < \infty,$

2. $K(s, t) = \overline{K(t, s)}$

genügt. Solche Integralgleichungen werden *Integralgleichungen mit symmetrischem Kern* genannt. Nach den Sätzen 1 und 2 aus 9.2.1. ist der entsprechende Fredholmsche Operator

$$A\varphi = \int_a^b K(s, t)\, \varphi(t)\, dt \tag{6}$$

kompakt und selbstadjungiert. Daher gilt für ihn der Satz von HILBERT-SCHMIDT (Satz 5 aus 4.6.). Wir benutzen nun diesen Satz bei der Suche nach Lösungen der Gleichung (5). Da für uns jetzt nur die Kompaktheit und Selbstadjungiertheit des Operators (6), nicht aber seine Integraldarstellung von Bedeutung sind, schreiben wir die Integralgleichung (5) in der symbolischen Form

$$\varphi = A\varphi + f. \tag{7}$$

Nach dem Satz von HILBERT-SCHMIDT existiert ein Orthogonalsystem von Eigenfunktionen $\{\psi_n\}$ zu den Eigenwerten $\{\lambda_n\}$ des Operators A, so daß jedes Element ξ aus L_2 in der Form

$$\xi = \sum_n a_n \psi_n + \xi' \qquad (A\xi' = 0)$$

darstellbar ist. Zerlegen wir die gegebene Funktion f auf diese Weise,

$$f = \sum_n b_n \psi_n + f' \qquad (Af' = 0), \tag{8}$$

und suchen die Lösung φ der Gleichung (7) in der Form

$$\varphi = \sum_n x_n \psi_n + \varphi' \qquad (A\varphi' = 0), \tag{9}$$

dann erhalten wir durch Einsetzen von (8) und (9) in (7)

$$\sum_n x_n \psi_n + \varphi' = \sum_n x_n \lambda_n \psi_n + \sum_n b_n \psi_n + f'.$$

Diese Gleichung ist genau dann erfüllt, wenn

$$\varphi' = f'$$

und

$$x_n(1 - \lambda_n) = b_n \qquad (n = 1, 2, \ldots),$$

d. h.

$$\varphi' = f',$$

$$x_n = \frac{b_n}{1 - \lambda_n} \quad \text{für} \quad \lambda_n \neq 1,$$

$$b_n = 0 \quad \text{für} \quad \lambda_n = 1$$

ist. Dabei stellt die letzte Gleichung eine notwendige und hinreichende Bedingung für die Lösbarkeit der Gleichung (7) dar, während die ersten beiden Gleichungen Bestimmungsgleichungen für die Lösung φ sind. Die Koordinaten x_n, für die der Eigenwert mit demselben Index gleich 1 ist, sind beliebig wählbar. Insgesamt ergibt sich also folgendes Resultat.

Satz 3. *Ist* 1 *kein Eigenwert des Operators* A, *dann hat die Gleichung* (7) *für beliebiges* f *eine eindeutig bestimmte Lösung. Ist* 1 *Eigenwert des Operators* A, *so ist die Gleichung* (7) *genau dann lösbar, wenn* f *orthogonal zu allen Eigenfunktionen des Operators* A *für den Eigenwert* 1 *ist. In diesem Fall hat die Gleichung* (7) *unendlich viele Lösungen.*

9.2.3. Die Fredholmschen Sätze für Integralgleichungen mit ausgeartetem Kern. Wir gehen nun zur Betrachtung Fredholmscher Integralgleichungen zweiter Art über, deren Kerne die Bedingung

$$\int_a^b \int_a^b |K(x, t)|^2 \, ds \, dt < \infty$$

erfüllen (daraus folgt die Kompaktheit des Operators), aber nicht symmetrisch sind. Zuerst betrachten wir eine Gleichung

$$\varphi(s) = \int_a^b K(s, t) \, \varphi(t) \, dt + f(s) \tag{10}$$

mit ausgeartetem Kern, d. h. einem Kern von der Form

$$K(s, t) = \sum_{i=1}^{n} P_i(s) \, Q_i(t), \tag{11}$$

wobei P_i, Q_i Funktionen aus L_2 sind. Ein Integraloperator mit einem ausgearteten Kern bildet jede Funktion $\varphi \in L_2$ auf die Summe

$$\sum_{i=1}^{n} P_i(s) \int_a^b Q_i(t) \, \varphi(t) \, dt,$$

d. h. auf ein Element des endlichdimensionalen Raumes ab, der durch die Funktion P_i $(i = 1, 2, \ldots, n)$ erzeugt wird. Wir können dabei annehmen, daß alle Funktionen $P_1, \ldots, P_n$ des Ausdrucks (11) voneinander linear unabhängig sind. Andernfalls würden wir jede Funktion P_i als Linearkombination von linear unabhängigen Funktionen darstellen, in (11) einsetzen und den Kern unter Beibehaltung seiner ausgearteten Struktur (aber kleineren Anzahl von Summanden) auf eine Form mit linear unabhängigen P_i bringen können. Denken wir uns eine entsprechende Reduktion auch für die Funktionen $Q_i(t)$ durchgeführt, so erhält der Kern (11) eine Darstellung, in der die P_i und Q_i voneinander linear unabhängig sind.

Für die Untersuchung der Lösbarkeit der Gleichung (10) mit einem ausgearteten Kern (11) können wir daher o. B. d. A. annehmen, daß die Funktionen $P_1, \ldots, P_n$ und auch die Funktionen $Q_1, \ldots, Q_n$ linear unabhängig sind. Setzen wir in (10) an Stelle von $K(s, t)$ die entsprechende Summe ein, so ergibt sich

$$\varphi(s) = \sum_{i=1}^{n} P_i(s) \int_a^b Q_i(t) \, \varphi(t) \, dt + f(s) \tag{12}$$

und daraus mit der Bezeichnung

$$\int_a^b Q_i(t) \, \varphi(t) = q_i$$

für die Funktion φ

$$\varphi(s) = \sum_{i=1}^{n} q_i P_i(s) + f(s).$$

Wird nun dieser Ausdruck für φ in (10) eingesetzt, so erhält die Integralgleichung die Form

$$\sum_{i=1}^{n} q_i P_i(s) + f(s) = \sum_{i=1}^{n} P_i(s) \int_a^b Q_i(t) \left[\sum_{j=1}^{n} q_j P_j(t) + f(t) \right] dt + f(s) \tag{13}$$

oder, wenn wir

$$\int_a^b Q_i(t) \, P_j(t) \, dt = a_{ij}, \qquad \int_a^b Q_i(t) \, f(t) \, dt = b_i$$

setzen,

$$\sum_{i=1}^{n} q_i P_i(s) = \sum_{i=1}^{n} P_i(s) \left[\sum_{j=1}^{n} a_{ij} q_j + b_i \right].$$

Da die Funktionen P_i nach Voraussetzung linear unabhängig sind, ist diese Gleichung nur erfüllt, wenn für die Koeffizienten

$$q_i = \sum_{i=1}^{n} a_{ij} q_j + b_i, \quad i = 1, 2, \ldots, n, \tag{14}$$

gilt, d. h., die Koeffizienten q_i müssen einem linearen Gleichungssystem genügen. Lösen wir dieses System und setzen die Lösungen q_i in den Ansatz

$$\varphi(s) = \sum_{i=1}^{n} q_i P_i(s) + f(s)$$

ein, dann erfüllt diese Funktion $\varphi(s)$ die Integralgleichung (10), da alle Schritte, die von der Gleichung (10) zum System (14) führten, auch in der umgekehrten Reihenfolge durchgeführt werden können. *Die Lösung einer Integralgleichung mit ausgeartetem Kern kann also auf die Lösung des ihr entsprechenden Systems* (14) *linearer algebraischer Gleichungen zurückgeführt werden.*

Aus den gut bekannten Existenz- und Eindeutigkeitsbedingungen für die Lösungen linearer Gleichungssysteme ergeben sich nun entsprechende Aussagen (Fredholmsche Sätze) für die Lösungen von Integralgleichungen mit ausgeartetem Kern. Wir formulieren im folgenden diese Sätze (auch im Hinblick auf die Resultate des nächsten Abschnittes) noch einmal.

I. Das System linearer algebraischer Gleichungen

$$Tx = y \quad \left(T = \|a_{ik}\|,\ x = (x_1, \ldots, x_n),\ y = (y_1, \ldots, y_n)\right)$$

ist genau dann lösbar, wenn der Vektor y zu allen Lösungen z des homogenen adjungierten Systems

$$T^* z = 0 \quad (T = \|\overline{a_{ki}}\|)$$

orthogonal ist.

II. Wenn die Determinante der Matrix T von 0 verschieden ist, dann hat die Gleichung $Tx = y$ für beliebiges y genau eine Lösung. Wenn die Determinante der Matrix T gleich 0 ist, dann hat die homogene Gleichung $Tx = 0$ nichttriviale Lösungen.

III. Da die Matrix T und die adjungierte Matrix T^* den gleichen Rang haben, besitzen die homogenen Systeme $Tx = 0$ und $T^*z = 0$ die gleiche Anzahl linear unabhängiger Lösungen.

Im nächsten Abschnitt werden wir zeigen, daß diese Sätze im wesentlichen auch für Integralgleichungen mit beliebigen (nichtausgearteten) Kernen richtig sind. Da

jedoch für nichtausgeartete Integraloperatoren solche Begriffe wie Rang, Matrix, Determinante keinen Sinn haben, werden wir die entsprechenden Sätze so formulieren, daß diese Begriffe darin nicht auftreten.

9.2.4. Die Fredholmschen Sätze für Integralgleichungen mit nichtausgeartetem Kern.

Wir betrachten wieder die Integralgleichung

$$\varphi(s) = \int_a^b K(s, t)\, \varphi(t)\, dt + f(s), \tag{15}$$

setzen aber jetzt nur voraus, daß der Kern die Hilbert-Schmidt-Bedingung

$$\int_a^b \int_a^b |K(s, t)|^2\, ds\, dt < \infty$$

erfüllt (daraus folgt die Kompaktheit des Operators). Wir setzen jetzt im Gegensatz zu den vorhergehenden Abschnitten weder voraus, daß der Kern ausgeartet, noch daß er symmetrisch ist. Von Interesse sind für uns Bedingungen für die Lösbarkeit der Integralgleichung (15) und Eigenschaften ihrer Lösungen. Bei den Untersuchungen wird wieder nur die Kompaktheit des der Gleichung (15) entsprechenden Operators und nicht seine Integraldarstellung wesentlich sein. Daher betrachten wir wie früher an Stelle von (15) die abstrakte Operatorgleichung

$$\varphi = A\varphi + f, \tag{16}$$

in der A ein beliebiger kompakter Operator im Hilbertraum H ist. Setzen wir $T = I - A$, wobei I der Einheitsoperator ist, so erhält (16) die Form

$$T\varphi = f. \tag{17}$$

Zusammen mit dieser Gleichung werden wir die homogene Gleichung

$$T\varphi_0 = 0 \tag{18}$$

und die adjungierten Gleichungen

$$T^*\psi = g, \tag{19}$$

$$T^*\psi_0 = 0 \tag{20}$$

($T^* = I - A^*$) betrachten. Zwischen den Lösbarkeitseigenschaften dieser vier Gleichungen besteht ein enger Zusammenhang, der durch die folgenden *Fredholmschen Sätze* aufgezeigt wird.

I. *Die inhomogene Gleichung $T\varphi = f$ ist genau für diejenigen f lösbar, die zu allen Lösungen der homogenen adjungierten Gleichung $T^*\psi_0 = 0$ orthogonal sind.*

II (*Fredholmsche Alternative*). *Es besitzt entweder die inhomogene Gleichung $T\varphi = f$ für beliebiges $f \in H$ eine eindeutig bestimmte Lösung oder die homogene Gleichung $T\varphi_0 = 0$ eine nichttriviale Lösung.*

III. *Die homogenen Gleichungen* (18) *und* (20) *haben stets ein und dieselbe endliche Anzahl linear unabhängiger Lösungen.*

Bevor wir diese Sätze beweisen, erwähnen wir noch, daß sie natürlich auch für die früher behandelten Integralgleichungen mit symmetrischen bzw. ausgearteten Kernen richtig sind. Unter den stärkeren Voraussetzungen von 9.2.2. fanden diese Sätze ihren Ausdruck im Satz 3 (der dritte Fredholmsche Satz ist in diesem Fall wegen $A = A^*$ trivial). Unter den Voraussetzungen von 9.2.3., die die Reduzierung der entsprechenden Integralgleichungen auf lineare algebraische Gleichungssysteme gestatten, wurden die Fredholmschen Sätze durch die angeführten Sätze über die Lösungen linearer Gleichungssysteme ausgedrückt.

Da jeder kompakte Operator als Grenzwert einer konvergenten Folge endlichdimensionaler Operatoren dargestellt werden kann und diese für einen kompakten Integraloperator als ausgeartete Integraloperatoren gewählt werden können, ergibt sich ein Beweis der Fredholmschen Sätze für Integralgleichungen mit nichtausgeartetem Kern durch Grenzübergang aus den entsprechenden Sätzen für Integralgleichungen mit ausgeartetem Kern. Wir gehen hier jedoch nicht diesen Weg.

Beweis der Fredholmschen Sätze. Ist B ein linearer stetiger Operator im Hilbertraum H, dann ist die Menge $\operatorname{Ker} B = \{x : x \in H, Bx = 0\}$, wie man leicht nachprüft, ein abgeschlossener linearer Unterraum des Raumes H. Der Wertevorrat $\operatorname{Im} B = \{y : y \in H, y = Bx\}$ des Operators B ist zwar auch eine lineare Menge, jedoch im allgemeinen nicht abgeschlossen. Wir zeigen nun, daß für den Operator $T = I - A$ der Wertevorrat eine abgeschlossene Menge ist.

Lemma 1. *Die Menge* $\operatorname{Im} T$ *ist abgeschlossen.*

Beweis. Es sei $y_n \in \operatorname{Im} T$ und $y_n \to y$. Dann existierten Vektoren $x_n \in H$, so daß

$$y_n = Tx_n = x_n - Ax_n \tag{21}$$

gilt. Diese Folge $\{x_n\}$ ist beschränkt und kann orthogonal zur Menge $\operatorname{Ker} T$ gewählt werden. Die Orthogonalität zu $\operatorname{Ker} T$ können wir sofort erreichen, wenn wir von den x_n in (21) nötigenfalls ihre Projektion auf $\operatorname{Ker} T$ subtrahieren. Die Beschränktheit der Folge $\{x_n\}$ zeigen wir indirekt. Aus der Annahme, daß für die Folge $\{x_n\}$ (oder eine Unterfolge) $\|x_n\| \to \infty$ gilt, und der Gültigkeit von (21) ergibt sich

$$\frac{x_n}{\|x_n\|} - A\left(\frac{x_n}{\|x_n\|}\right) \to 0.$$

Da der Operator A kompakt ist, d. h. $\left\{A\left(\frac{x_n}{\|x_n\|}\right)\right\}$ eine konvergente Teilfolge enthält, konvergiert auch eine Teilfolge von $\left\{\frac{x_n}{\|x_n\|}\right\}$. Bezeichnen wir ihren Grenzwert mit z, so ist $\|z\| = 1$ und $Tz = 0$, d. h. $z \in \operatorname{Ker} T$. Andererseits muß z orthogonal zur Menge $\operatorname{Ker} T$ sein, da alle x_n orthogonal zu $\operatorname{Ker} T$ gewählt wurden. Aus diesem Widerspruch folgt nun, daß die Folge $\{x_n\}$ beschränkt ist. Auf Grund der Kompaktheit des Opera-

tors A enthält daher die Folge $\{Ax_n\}$ eine konvergente Teilfolge, was nach (21) auch die Existenz einer konvergenten Teilfolge der Folge $\{x_n\}$ nach sich zieht. Ist x der Grenzwert dieser Teilfolge, dann folgt aus (21), daß $y = Tx$ ist. Damit ist das Lemma bewiesen.

Lemma 2. *Der Raum H ist die direkte orthogonale Summe der abgeschlossenen Unterräume* Ker T *und* Im T^*, *d. h.*

$$\operatorname{Ker} T \oplus \operatorname{Im} T^* = H. \tag{22}$$

Entsprechend gilt auch

$$\operatorname{Ker} T^* \oplus \operatorname{Im} T = H. \tag{23}$$

Beweis. Wir wissen bereits, daß beide Unterräume auf der linken Seite der Gleichung (22) abgeschlossen sind. Außerdem sind sie offensichtlich orthogonal, denn für $h \in \operatorname{Ker} T$ gilt $(h, T^*x) = (Th, x) = 0$ für alle $x \in H$. Damit ist nur noch zu zeigen, daß es keinen von 0 verschiedenen Vektor in H gibt, der gleichzeitig zu Ker T und Im T^* orthogonal ist. Nehmen wir an, z sei ein solcher Vektor, so folgt aus $z \perp \operatorname{Im} T^*$ für ein beliebiges $x \in H$

$$0 = (z, T^*x) = (Tz, x),$$

d. h. $z \in \operatorname{Ker} T$, was der Voraussetzung über z widerspricht. Damit ist das Lemma bewiesen.

Aus Lemma 2 folgt unmittelbar der erste Fredholmsche Satz, denn $f \perp \operatorname{Ker} T^*$ gilt genau in dem Fall, wenn $f \in \operatorname{Im} T$ ist, d. h., wenn ein Vektor φ mit der Eigenschaft $T\varphi = f$ existiert.

Setzen wir $\operatorname{Im}(T^k) = H^k$ $(k = 1, 2, \ldots)$, d. h. $\operatorname{Im} T = H^1$, $T(H^k) = H^{k+1}$, so sind die H^k nach Lemma 1 abgeschlossene Unterräume, die offensichtlich in der Beziehung

$$H \supset H^1 \supset H^2 \supset \cdots \tag{24}$$

zueinander stehen.

Lemma 3. *Es gibt eine Zahl j, so daß $H^{k+1} = H^k$ für alle $k \geqq j$ ist.*

Beweis. Wir nehmen an, daß es keine solche Zahl j gibt. In diesem Fall sind alle H^k verschieden, und man kann eine orthonormierte Folge $\{x_k\}$ konstruieren, so daß $x_k \in H^k$ und $x_k \perp H^{k+1}$ ist. Für $l > k$ gilt dann

$$Ax_l - Ax_k = -x_k + (x_l + Tx_k - Tx_l),$$

wobei $(x_l + Tx_k - Tx_l) \in H^{k+1}$ ist. Daraus folgt $\|Ax_l - Ax_k\| \geqq 1$. Das bedeutet, daß aus $\{Ax_k\}$ keine konvergente Teilfolge ausgewählt werden kann. Dieser Widerspruch zur Kompaktheit des Operators A zeigt, daß die Annahme falsch ist, was zu zeigen war.

Lemma 4. *Ist* $\operatorname{Ker} T = \{0\}$, *dann gilt* $\operatorname{Im} T = H$.

Beweis. $\operatorname{Ker} T = \{0\}$ bedeutet, daß der Operator T eineindeutig ist. Wäre jetzt $\operatorname{Im} T \neq H$, dann bestünde die Kette (24) aus lauter verschiedenen Unterräumen, was Lemma 3 widerspricht. Daher gilt $\operatorname{Im} T = H$, wenn $\operatorname{Ker} T = \{0\}$ ist. Analog ergibt sich $\operatorname{Im} T^* = H$ im Fall $\operatorname{Ker} T^* = 0$.

Lemma 5. *Ist* $\operatorname{Im} T = H$, *dann gilt* $\operatorname{Ker} T = \{0\}$.

Beweis. Aus $\operatorname{Im} T = H$ folgt $\operatorname{Ker} T^* = \{0\}$ nach Lemma 2, daraus $\operatorname{Im} T^* = H$ nach Lemma 4 und hieraus $\operatorname{Ker} T = \{0\}$ nach Lemma 2.

Die Aussage der Lemmata 4 und 5 stellt die Fredholmsche Alternative (zweiter Fredholmscher Satz) dar.

Wir beweisen nun den dritten Fredholmschen Satz. Um zu zeigen, daß die homogenen Gleichungen nur endlich viele linear unabhängige Lösungen besitzen, nehmen wir im Sinne eines indirekten Beweises zunächst an, daß der Unterraum $\operatorname{Ker} T$ unendlichdimensional ist. In diesem Fall existiert in $\operatorname{Ker} T$ ein unendliches orthonormiertes System $\{x_k\}$, für das $Ax_k = x_k$ gilt. Ist $l \neq k$, so ergibt sich $\|Ax_k - Ax_l\| = \sqrt{2}$. Daher kann aus der Folge $\{Ax_k\}$ keine konvergente Unterfolge ausgewählt werden. Das steht im Widerspruch zur Kompaktheit des Operators A. Der Widerspruch zeigt, daß die Annahme falsch war, also $\operatorname{Ker} T$ und entsprechend $\operatorname{Ker} T^*$ endlichdimensionale Unterräume sind.

Ist die Dimension von $\operatorname{Ker} T$ gleich μ und die von $\operatorname{Ker} T^*$ gleich ν, so bleibt nun noch $\mu = \nu$ zu zeigen. Dazu nehmen wir zunächst an, daß $\mu < \nu$ ist, und leiten daraus einen Widerspruch ab. Es sei $\{\varphi_1, \ldots, \varphi_\mu\}$ eine orthonormierte Basis in $\operatorname{Ker} T$ und $\{\psi_1, \ldots, \psi_\nu\}$ eine orthonormierte Basis in $\operatorname{Ker} T^*$. Setzen wir

$$Sx = Tx + \sum_{j=1}^{\mu} (x, \varphi_j)\, \psi_j,$$

so ist dieser Operator S offensichtlich kompakt, da er sich von dem kompakten Operator T nur durch einen endlichdimensionalen Operator unterscheidet.

Aus der Gleichung $Sx = 0$, d. h.

$$Tx + \sum_{j=1}^{\mu} (x, \varphi_j)\, \psi_j = 0, \tag{25}$$

ergibt sich

$$Tx = 0$$

und

$$(x, \varphi_j) = 0 \qquad \text{für } 1 \leqq j \leqq \mu,$$

da nach Lemma 2 alle Vektoren ψ_j zu allen Vektoren der Form Tx orthogonal sind. Die Lösung x der Gleichung $Sx = 0$ ist also einerseits als Element aus $\operatorname{Ker} T$ eine Linearkombination der Vektoren φ_j und andererseits zu allen φ_j orthogonal, also stets gleich 0. Nach dem zweiten Fredholmschen Satz (angewandt auf den kom-

pakten Operator S) besitzt nun die Gleichung $Sy = \psi_{\mu+1}$, d. h.

$$Ty + \sum_{j=1}^{\mu} (y, \varphi_j)\, \psi_j = \psi_{\mu+1},$$

eine eindeutig bestimmte Lösung $y \in H$. Multiplizieren wir diese Gleichung skalar mit $\psi_{\mu+1}$, so ergibt sich rechts 1 und links 0 wegen $Ty \in \operatorname{Im} T$ und $\operatorname{Im} T \perp \operatorname{Ker} T^*$. Dieser Widerspruch entsteht aus der Annahme $\mu < \nu$. Daher kann nur $\mu \geqq \nu$ richtig sein. Geht man von der Annahme $\mu > \nu$ aus und führt die obigen Betrachtungen für den Operator T^* (an Stelle von T) durch, so erhält man wie oben einen Widerspruch. Daher kann nur $\mu = \nu$ richtig sein, was zu zeigen war.

Bemerkungen

1. In den Fredholmschen Sätzen werden Aussagen über die Umkehrbarkeit des Operators $A - I$ gemacht, insbesondere wird festgestellt, daß $\lambda = 1$ entweder ein regulärer Punkt oder ein Eigenwert endlicher Vielfachheit für den Operator A ist. Alle diese Aussagen bleiben offensichtlich richtig, wenn statt $A - I$ der Operator $A - \lambda I$ mit $\lambda \neq 0$ betrachtet wird. Daraus ergibt sich der Satz: *Jeder von 0 verschiedene Punkt des Spektrums eines kompakten Operators A ist ein Eigenwert endlicher Vielfachheit.* Nach Satz 4 aus 4.6. ist die Menge dieser Eigenwerte höchstens abzählbar. Der Punkt 0 gehört immer zum Spektrum eines kompakten Operators A in einem unendlichdimensionalen Raum (vgl. Satz 2 aus 4.6.), ist jedoch nicht notwendig ein Eigenwert von A. Kompakte Operatoren, deren Spektrum nur aus dem Punkt 0 besteht, werden (abstrakte) *Volterrasche Operatoren* genannt.

2. Wir haben die Fredholmschen Sätze für Gleichungen der Form $\varphi = A\varphi + f$ (A ein kompakter Operator) in einem Hilbertraum H bewiesen. Alle Sätze sind ohne wesentliche Änderungen auch für Gleichungen dieser Art in einem beliebigen Banachraum E richtig. Dabei ist die adjungierte Gleichung $\psi = A^*\psi + g$ als Gleichung im adjungierten Raum E^* und die Bedingung der Orthogonalität $(f, \psi_0) = 0$ als Bedingung $\psi_0(f) = 0$ über die Werte der Funktionale $\psi_0 \in \operatorname{Ker} T^* \subset E^*$ für das Element $f \in E$ zu verstehen. Eine Darstellung der Fredholmschen Sätze für Gleichungen in einem Banachraum ist z. B. in [39] enthalten.

9.2.5. Volterrasche Integralgleichungen. Die Integralgleichung

$$\varphi(s) = \int_a^s K(s, t)\, \varphi(t)\, dt + f(s), \tag{26}$$

in der $K(s, t)$ eine beschränkte meßbare Funktion ist, wird *Volterrasche Integralgleichung zweiter Art* genannt. Diese Gleichung ist ein Spezialfall der Fredholmschen Integralgleichung (man setzt $K(s, t) = 0$ für $t > s$). Damit gelten für sie die Fredholmschen Sätze, die jedoch jetzt zu der folgenden Aussage verschärft werden können: *Die Volterrasche Integralgleichung* (26) *besitzt für eine beliebige Funktion $f \in L_2[a, b]$ stets eine eindeutig bestimmte Lösung.*

Wiederholen wir nämlich für den Operator

$$A\varphi = \int_a^s K(s, t)\, \varphi(t)\, dt$$

die Überlegungen aus 2.4.4., so ergibt sich, daß eine Potenz dieses Operators eine kontrahierende Abbildung ist und demzufolge die homogene Gleichung nur die triviale Lösung besitzt; ausgehend von dieser Tatsache lassen sich dann die Fredholmschen Sätze für Volterrasche Integralgleichungen zweiter Art in der obigen Form zusammenfassen.

Aufgabe. Auf einem Intervall $[a, b]$ sei eine Fredholmsche Integralgleichung zweiter Art mit einem stetigen reellen Kern gegeben. Man zeige, daß die Fredholmschen Sätze auch dann gelten, wenn diese Gleichung im Raum $C[a, b]$ der stetigen Funktionen betrachtet wird. Dabei ist als „adjungierte Gleichung" die Integralgleichung mit dem transponierten Kern $K(t, s)$ und die Orthogonalität wie im reellen Raum L_2 zu verstehen.

9.2.6. Integralgleichungen erster Art. Eine Gleichung

$$A\varphi = f, \tag{27}$$

in der A ein kompakter Operator in einem Hilbertraum H, f eine gegebene und φ eine gesuchte Funktion aus H ist, heißt abstrakte *Fredholmsche Gleichung erster Art.*

Die Aufgabe, eine solche Gleichung zu lösen, ist im allgemeinen schwieriger als bei einer Gleichung zweiter Art. Wir betrachten zunächst als einfaches Beispiel im Raum $L_2[a, b]$ die Integralgleichung

$$f(s) = \int_a^s \varphi(t)\, dt, \tag{28}$$

d. h. eine Gleichung mit dem Kern

$$K(s, t) = \begin{cases} 1 \text{ für } t \leqq s, \\ 0 \text{ für } t > s. \end{cases}$$

Sie besitzt offensichtlich die Lösung $\varphi(s) = f'(s)$, wenn f absolut stetig ist und $f'(s)$ zu L_2 gehört, aber keine Lösung, wenn f diese Voraussetzungen nicht erfüllt.

Wir zeigen nun, daß auch im allgemeinen Fall die Gleichung (27) nicht für jedes $f \in H$ lösbar ist. Die Existenz einer Lösung von $A\varphi = f$ für beliebiges $f \in H$ würde bedeuten, daß der Operator A den Raum H auf ganz H abbildet. Das ist jedoch bei einem kompakten Operator A nicht der Fall. Denn stellt man H als Vereinigung abzählbar vieler Kugeln S_n (z. B. Kugeln um den Nullpunkt mit dem Radius $1, 2, \ldots, n, \ldots$) dar, so erscheint AH als abzählbare Vereinigung der präkompakten Mengen AS_n. Da präkompakte Mengen im endlichdimensionalen Hilbertraum H nirgends dicht sind, würde aus $H = AH$ folgen, daß H als abzählbare Vereinigung

nirgends dichter Mengen darstellbar ist, was nach Satz 2 aus 2.3. nicht möglich ist. Daher gilt $H \neq AH$ oder, mit anderen Worten, der Satz: *Ist A ein beliebiger kompakter Operator A in H, so ist die Gleichung $A\varphi = f$ nicht für alle $f \in H$ lösbar.*

Ein anderer wesentlicher Unterschied zwischen den Gleichungen $(A - I)\,\varphi = f$ und $A\varphi = f$ mit einem kompakten Operator A im Hilbertraum H besteht darin, daß der Umkehroperator $(A - I)^{-1}$ beschränkt, der Operator A^{-1} aber unbeschränkt ist. Sind f_1 und f_2 zwei nahe beieinanderliegende Elemente aus H, für die die Gleichungen

$$A\varphi_1 = f_1, \qquad A\varphi_2 = f_2$$

lösbar sind, dann können also die Lösungen $\varphi_1 = A^{-1}f_1$ und $\varphi_2 = A^{-1}f_2$ sehr stark voneinander abweichen, oder, anders ausgedrückt, ein beliebig kleiner Fehler der rechten Seite von Gleichung (27) kann zu einem beliebig großen Fehler der Lösung führen. Aufgaben, in denen eine kleine Änderung der vorgegebenen Größen zu kleinen Änderungen der Lösung führt, heißen *korrekt* gestellte Aufgaben (dabei kann „kleine Änderung" in verschiedenen Aufgaben verschieden zu verstehen sein). Die Bestimmung der Lösung einer Integralgleichung erster Art ist (im Unterschied zur Bestimmung der Lösung einer Integralgleichung zweiter Art) eine nicht korrekt gestellte Aufgabe. In der letzten Zeit wurden nicht korrekt gestellte Aufgaben verschiedener Art betrachtet und Methoden für ihre Regularisierung (d. h. für ihre Zurückführung auf in diesem oder jenem Sinn korrekt gestellte Aufgaben) entwickelt. Eine Darlegung dieser Fragen ist jedoch im Rahmen dieses Buches nicht möglich

9.3. Integralgleichungen, die einen Parameter enthalten. Die Fredholmsche Methode

9.3.1. Das Spektrum eines kompakten Operators im Hilbertraum. Wir betrachten die Gleichung

$$\varphi = \lambda A\varphi + f$$

oder, anders geschrieben,

$$(I - \lambda A)\,\varphi = f, \tag{1}$$

wobei A ein kompakter Operator im Hilbertraum H und λ ein Zahlenparameter ist.

Nach der Fredholmschen Alternative gibt es bezüglich der Lösbarkeit von (1) nur zwei einander ausschließende Möglichkeiten:

1. Die Gleichung (1) besitzt bei vorgegebenem λ für jedes $f \in H$ eine eindeutig bestimmte Lösung.

2. Die homogene Gleichung $\varphi = \lambda A\varphi$ besitzt eine nichttriviale Lösung.

Im ersten Fall bildet der Operator $I - \lambda A$ den Raum H eineindeutig auf ganz H ab. Daraus folgt, daß dann auf ganz H ein beschränkter Umkehroperator $(I - \lambda A)^{-1}$ existiert.

Gleichbedeutend damit ist die Aussage, daß $\left(A - \frac{1}{\lambda} I\right)^{-1}$ auf ganz H definiert und beschränkt ist, d. h. $\frac{1}{\lambda}$ nicht zum Spektrum des Operators A gehört.

Liegt der zweite Fall vor, so existiert ein von 0 verschiedenes Element $\varphi_\lambda \in H$ mit der Eigenschaft

$$\varphi_\lambda = \lambda A \varphi_\lambda \quad \text{oder} \quad A\varphi_\lambda = \frac{1}{\lambda} \varphi_\lambda,$$

d. h., $\frac{1}{\lambda}$ ist ein Eigenwert des Operators A.

Damit haben wir folgendes Resultat hergeleitet: *Jede von 0 verschiedene Zahl $\mu = \frac{1}{\lambda}$ ist entweder regulär oder ein Eigenwert für den kompakten Operator A.*

Verbinden wir dieses Resultat mit der Aussage des Satzes 4 in 4.6., so ergibt sich folgende Beschreibung des Spektrums eines kompakten Operators A im Hilbertraum H. Das Spektrum eines beliebigen kompakten Operators A in H besteht aus endlich oder abzählbar vielen von 0 verschiedenen Eigenwerten $\mu_1, \mu_2, \ldots, \mu_n, \ldots$ endlicher Vielfachheit und dem Punkt 0[1]). Die Folge $\{\mu_n\}$ häuft sich höchstens an der Stelle 0. Der Punkt $\mu = 0$ selbst kann ein Eigenwert endlicher oder unendlicher Vielfachheit oder auch nur Häufungspunkt der Menge der Eigenwerte sein.

Für einen Volterraschen Integraloperator B enthält das Spektrum nur den Punkt 0, da die Gleichung

$$\varphi = \lambda B \varphi + f$$

für jedes $f \in L_2$ eindeutig lösbar ist (vgl. 9.2.5.). Daraus resultiert auch die Bezeichnung „abstrakter Volterrascher Operator" für einen kompakten Operator A, dessen Spektrum nur aus dem Punkt 0 besteht.

9.3.2. Der Potenzreihenansatz für die Lösung einer Integralgleichung. Fredholmsche Determinanten. Die Lösung der Gleichung

$$(I - \lambda A)\, \varphi = f$$

kann man formal durch

$$\varphi = (I - \lambda A)^{-1} f \tag{2}$$

beschreiben. Für $\|\lambda A\| < 1$, d. h. $|\lambda| < \frac{1}{\|A\|}$, ist der Operator $(I - \lambda A)^{-1}$ stets auf ganz H definiert, beschränkt und als Summe

$$(I - \lambda A)^{-1} = I + \lambda A + \lambda^2 A^2 + \cdots + \lambda^n A^n + \cdots$$

---

[1]) Der Punkt $\mu = 0$ gehört notwendig immer zum Spektrum des Operators A, da A^{-1} in H nicht beschränkt ist.

einer Potenzreihe darstellbar, deren Konvergenz (in der Operatornorm) durch die Bedingung $|\lambda| < \frac{1}{\|A\|}$ gesichert wird. Daher kann man in diesem Fall die Lösung (2) von (1) in der Form

$$y = f + \lambda A f + \lambda^2 A^2 f + \cdots + \lambda^n A^n f + \cdots \tag{3}$$

schreiben. Das gleiche Resultat ergibt sich, wenn man für die Lösung von (1) einen Potenzreihenansatz

$$\varphi_\lambda = \varphi_0 + \lambda \varphi_1 + \cdots + \lambda^n \varphi_n + \cdots$$

macht, wobei φ_n nicht von λ abhängt. Wird dieser Ansatz in die linke und die rechte Seite der Gleichung $\varphi = \lambda A \varphi + f$ eingesetzt, so liefert der Vergleich der Koeffizienten von gleichen Potenzen λ^k

$$\varphi_0 = f, \quad \varphi_1 = A\varphi_0 = Af, \ldots, \quad \varphi_n = A\varphi_{n-1} = A^n f, \ldots,$$

d. h. die Reihe (3).

Ist A ein Hilbert-Schmidt-Integraloperator (d. h., ist der zugehörige Kern $K(s, t)$ quadratisch integrierbar), dann kann der Operator $(I - \lambda A)^{-1}$ für genügend kleine Parameterwerte λ als Summe $I + \lambda\Gamma(\lambda)$ des Einheitsoperators I und eines von λ abhängigen Hilbert-Schmidt-Integraloperators $\lambda\Gamma(\lambda)$ dargestellt werden. Bevor wir diese Aussage beweisen, berechnen wir vorbereitend die Kerne, die zu den Operatoren A^2, A^3 usw. gehören. Es seien zwei Integraloperatoren

$$A\varphi = \int_a^b K(s, t)\, \varphi(t)\, dt,$$

$$B\varphi = \int_a^b Q(s, t)\, \varphi(t)\, dt$$

gegeben, wobei für die Kerne dieser Operatoren

$$\int_a^b \int_a^b |K(s, t)|^2\, ds\, dt = k^2 < \infty,$$

$$\int_a^b \int_a^b |Q(s, t)|^2\, ds\, dt = q^2 < \infty$$

gilt. Gesucht ist der Kern des Operators AB. Nach dem Satz von FUBINI ist

$$AB\varphi = \int_a^b \left\{ K(s, u) \int_a^b Q(u, t)\, \varphi(t)\, dt \right\} du$$

$$= \int_a^b \left\{ \int_a^b K(s, u)\, Q(u, t)\, du \right\} \varphi(t)\, dt,$$

da die Funktionen $K(s, u)\,\varphi(t)$ und $Q(u, t)$ für $(u, t) \in [a, b] \times [a, b]$ quadratisch integrierbar und folglich das Produkt

$$K(s, u)\, Q(u, t)\, \varphi(t)$$

über dieselbe Menge integrierbar ist.

Setzen wir

$$R(s, t) = \int_a^b K(s, u)\, Q(u, t)\, du, \tag{4}$$

dann ergibt sich nach der Schwarzschen Ungleichung

$$|R(s, t)|^2 \leqq \int_a^b |K(s, u)|^2\, du \cdot \int_a^b |Q(u, t)|^2\, du$$

und daraus

$$\int_a^b \int_a^b |R(s, t)|^2\, ds\, dt \leqq k^2 q^2,$$

d. h., das Produkt zweier Hilbert-Schmidt-Integraloperatoren ist wieder ein Integraloperator vom gleichen Typ, dessen Kern durch (4) definiert wird. Insbesondere ist also A^2 ein Integraloperator mit dem Kern

$$K_2(s, t) = \int_a^b K(s, u)\, K(u, t)\, du,$$

der die Bedingung

$$\int_a^b \int_a^b |K_2(s, t)|^2\, ds\, dt \leqq \left[\int_a^b \int_a^b |K(s, t)|^2\, ds\, dt\right]^2 = k^4$$

erfüllt. Daraus folgt $\|A^2\| \leqq k^2$.

Analog findet man, daß jeder Operator A^n ein Integraloperator mit dem Kern

$$K_n(s, t) = \int_a^b K_{n-1}(s, u)\, K(u, t)\, du \qquad (n = 2, 3, \ldots),$$

ist, der die Bedingung

$$\int_a^b \int_a^b |K_n(s, t)|^2\, ds\, dt \leqq k^{2n} \tag{5}$$

erfüllt. Die Kerne $K_n(s, t)$ heißen *iterierte Kerne*.

Für $|\lambda| < \dfrac{1}{k}$ konvergiert nun die Reihe

$$K(s, t) + \lambda K_2(s, t) + \cdots + \lambda^{n-1} K_n(s, t) + \cdots$$

auf Grund der Abschätzung (5) im Raum $L_2([a, b] \times [a, b])$ gegen eine quadratisch integrierbare Funktion $\Gamma(s, t; \lambda)$. Der Integraloperator $\Gamma(\lambda)$ mit dem Kern $\Gamma(s, t; \lambda)$ ist dann gleich der Summe der konvergenten Reihe

$$A + \lambda A^2 + \cdots + \lambda^{n-1} A^n + \cdots. \tag{6}$$

Als Grenzwert kompakter Operatoren ist er ebenfalls kompakt. Multiplizieren wir die Summe (6) mit λ und addieren den Einheitsoperator hinzu, so erhalten wir gerade den Operator $(I - \lambda A)^{-1}$, d. h., der Operator $(I - \lambda A)^{-1}$ ist die Summe aus dem Einheitsoperator I und dem kompakten Operator $\lambda\Gamma(\lambda)$ mit dem Kern

$$\lambda\Gamma(s, t; \lambda) = \sum_{n=1}^{\infty} \lambda^n K_n(s, t).$$

was zu zeigen war.

Die Bedingung $|\lambda| < \frac{1}{k}$ ist dabei hinreichend für die Konvergenz der Reihe (6), aber keineswegs notwendig. In einigen Fällen erweist sich diese Reihe sogar als für alle λ konvergent, z. B. wenn A ein Volterrascher Operator ist. Gilt nämlich für den Kern $K(s, t)$ eines solchen Operators

$$|K(s, t)| \leqq M,$$

dann genügen die iterierten Kerne, wie man durch direkte Rechnung zeigt, der Abschätzung

$$|K_n(s, t)| \leqq \frac{M^n (b - a)^{n-1}}{(n - 1)!}.$$

Daraus folgt, daß die Reihe (6) in diesem Fall für jedes λ konvergiert.

Im allgemeinen Fall besitzt die Potenzreihe (6) jedoch einen endlichen Konvergenzradius, d. h., die Frage nach der Konstruktion der Lösung φ von $\varphi = \lambda A\varphi + f$ in allen regulären Fällen $\left(\frac{1}{\lambda} \text{ kein Eigenwert von } A\right)$ wird durch (6) nicht vollständig beantwortet. Für Integralgleichungen ist eine vollständige Antwort möglich. FREDHOLM fand sie zunächst für Integraloperatoren A mit einem beschränkten und stetigen Kern $K(s,t)$ in der folgenden Weise. Es sei $K\begin{pmatrix} s_1 \dots s_n \\ t_1 \dots t_n \end{pmatrix}$ die Determinante

$$\begin{vmatrix} K(s_1, t_1) \dots K(s_1, t_n) \\ \dots\dots\dots\dots\dots\dots \\ K(s_n, t_1) \dots K(s_n, t_n) \end{vmatrix},$$

und $D(\lambda)$ und $D(s, t; \lambda)$ seien die sogenannten *Fredholmschen Determinanten*, die durch die Formeln

$$D(\lambda) = 1 - \lambda \int_a^b K\begin{pmatrix}\xi_1\\ \xi_1\end{pmatrix} d\xi_1 + \frac{\lambda^2}{2}\int_a^b\int_a^b K\begin{pmatrix}\xi_1 & \xi_2\\ \xi_1 & \xi_2\end{pmatrix} d\xi_1\, d\xi_2 + \cdots$$

$$+ (-1)^n \frac{\lambda^n}{n!}\int_a^b \cdots \int_a^b K\begin{pmatrix}\xi_1 & \ldots & \xi_n\\ \xi_1 & \ldots & \xi_n\end{pmatrix} d\xi_1 \cdots d\xi_n + \cdots, \tag{7}$$

$$D(s, t; \lambda) = K\begin{pmatrix}s\\ t\end{pmatrix} - \lambda \int_a^b K\begin{pmatrix}s & \xi_1\\ t & \xi_1\end{pmatrix} d\xi_1 + \frac{\lambda^2}{2}\int_a^b\int_a^b K\begin{pmatrix}s & \xi_1 & \xi_2\\ t & \xi_1 & \xi_2\end{pmatrix} d\xi_1\, d\xi_2 + \cdots$$

$$+ (-1)^n \frac{\lambda^n}{n!}\int_a^b \cdots \int_a^b K\begin{pmatrix}s & \xi_1 & \ldots & \xi_n\\ t & \xi_1 & \ldots & \xi_n\end{pmatrix} d\xi_1 \cdots d\xi_n + \cdots \tag{8}$$

definiert werden. Dann ist die Lösung der Integralgleichung

$$\varphi(s) = \lambda \int_a^b K(s, t)\, \varphi(t)\, dt + f(s)$$

für alle diejenigen λ, für die $\frac{1}{\lambda}$ kein Eigenwert des von $K(s, t)$ erzeugten Integraloperators A ist, in der Form

$$\varphi(s) = f(s) + \lambda \int_a^b \frac{D(s, t; \lambda)}{D(\lambda)} f(t)\, dt \tag{9}$$

darstellbar. Das heißt, wenn der Umkehroperator $(I - \lambda A)^{-1}$ existiert, ist er durch (9) mit dem lösenden Kern

$$\Gamma(s, t; \lambda) = \lambda \frac{D(s, t; \lambda)}{D(\lambda)}$$

darstellbar. Dabei sind $D(\lambda)$ und $D(s, t; \lambda)$ ganze analytische Funktionen des Parameters λ, und $D(\lambda)$ verschwindet dann und nur dann, wenn $\frac{1}{\lambda}$ ein Eigenwert des Integraloperators A ist.

Carleman zeigte 1921, daß die von Fredholm (unter der Voraussetzung der Stetigkeit des Kernes $K(s, t)$) gefundenen Formeln (7), (8) und (9) auch im Fall eines beliebigen quadratisch integrierbaren Kerns die Lösung der entsprechenden Integralgleichung beschreiben.[1])

[1]) Vgl. T. Carleman, Zur Theorie der Integralgleichungen, Math. Z. **9** (1921), 196–217, oder auch F. Smithies, The Fredholm theory of integral equations, Duke Math. J. **8** (1941), 107–130. Ein Beweis für die Formel (9) wird auch in den Büchern [43] und [64] angegeben.

10. Elemente der Differentialrechnung in linearen Räumen

In den bisher betrachteten Fragen der Funktionalanalysis spielte der Begriff des linearen Funktionals und des linearen Operators eine grundlegende Rolle. Es gibt jedoch auch Fragestellungen der Funktionalanalysis mit nichtlinearem Charakter. Sie machen es notwendig, zusammen mit der „linearen" auch eine „nichtlineare" Funktionalanalysis zu entwickeln, d. h. nichtlineare Funktionale und nichtlineare Operatoren in unendlichdimensionalen Räumen zu untersuchen. Zur nichtlinearen Funktionalanalysis gehört z. B. auch so ein klassisches Gebiet der Mathematik wie die Variationsrechnung, deren Grundlagen schon im 17. und 18. Jahrhundert von den Brüdern Bernoulli, von Euler und von Legendre geschaffen wurden. Dennoch ist die nichtlineare Funktionalanalysis als Ganzes noch ein verhältnismäßig junges Teilgebiet der Mathematik und von ihrer Vollendung zur Zeit noch weit entfernt. Wir stellen in diesem Kapitel nur einige Grundbegriffe der nichtlinearen Funktionalanalysis und einige Anwendungen dieser Begriffe dar.

10.1. Differentiation in linearen Räumen

10.1.1. Das starke Differential (Fréchetsches Differential). Es seien X und Y zwei lineare normierte Räume, und F sei eine Abbildung, die eine offene Untermenge O von X in eine Untermenge von Y abbildet. Wir nennen die Abbildung F *differenzierbar* in einem festen Punkt $x \in O$, wenn ein beschränkter linearer Operator $L_x \in \mathscr{L}(X, Y)$ existiert, so daß

$$F(x + h) - F(x) = L_x h + \alpha(x, h) \tag{1}$$

und für die Abbildung $\alpha(x, h)$

$$\frac{\|\alpha(x, h)\|}{\|h\|} \to 0 \quad \text{für} \quad \|h\| \to 0 \tag{2}$$

gilt. Der Ausdruck $L_x h$ (der offensichtlich für jedes h ein Element des Raumes Y darstellt) heißt *starkes Differential* (oder *Fréchetsches Differential*) der Abbildung F im Punkt x. Der lineare Operator L_x selbst heißt *Ableitung*, genauer *starke Ableitung*, der Abbildung F im Punkt x und wird mit dem Symbol $F'(x)$ bezeichnet.

Wenn die Abbildung F im Punkt x differenzierbar ist, dann ist die entsprechende Ableitung eindeutig bestimmt. Aus

$$F(x+h) - F(x) = L_x^{(1)}h + \alpha_1(x,h) = L_x^{(2)}h + \alpha_2(x,h)$$

folgt nämlich zunächst

$$L_x^{(1)}h - L_x^{(2)}h = \alpha_2(x,h) - \alpha_1(x,h)$$

und daraus nach (2)

$$\frac{\|L_x^{(1)}h - L_x^{(2)}h\|}{\|h\|} \to 0 \quad \text{für} \quad \|h\| \to 0. \tag{3}$$

Wäre nun $L_x^{(1)} \neq L_x^{(2)}$, d. h. für ein Element h

$$\frac{\|L_x^{(1)}h - L_x^{(2)}h\|}{\|h\|} = \lambda \neq 0,$$

so würde aus der Linearität der Operatoren

$$\frac{\|L_x^{(1)}(\varepsilon h) - L_x^{(2)}(\varepsilon h)\|}{\|\varepsilon h\|} = \lambda$$

folgen, was im Widerspruch zu (3) steht.

In gleicher Weise ergeben sich unmittelbar aus der Definition der Ableitung folgende Eigenschaften.

1. *Ist* $F(x) = y_0 = \text{const}$, *dann ist* $F'(x) \equiv 0$ (*d. h.,* $F'(x)$ *ist in diesem Fall der Nulloperator*).

2. *Die Ableitung einer stetigen linearen Abbildung stimmt mit der Abbildung selbst überein.*

Nach Definition ist nämlich

$$L(x+h) - L(x) = L(h).$$

Nicht ganz so offensichtlich ist die folgende Eigenschaft.

3. (*Differentiation einer zusammengesetzten Abbildung.*) *Es seien* X, Y, Z *drei normierte Räume,* $U(x_0)$ *sei eine Umgebung des Punktes* $x_0 \in X$, F *eine stetige Abbildung von* $U(x_0)$ *in* Y, $y_0 = F(x_0)$, $V(y_0)$ *eine Umgebung des Punktes* $y_0 \in Y$ *und* G *eine stetige Abbildung von* $V(y_0)$ *in* Z. *Ist* F *differenzierbar im Punkt* x_0 *und* G *differenzierbar im Punkt* y_0, *dann ist die zusammengesetzte Abbildung* $H = GF$ (*die auf einer Umgebung von* x_0 *definiert und stetig ist*) *differenzierbar im Punkt* x_0, *und es gilt*

$$H'(x_0) = G'(y_0)\, F'(x_0). \tag{4}$$

Um diese Aussage zu beweisen, setzen wir die Voraussetzungen[1])

$$F(x_0 + \xi) = F(x_0) + F'(x_0)\,\xi + o_1(\xi)$$

und

$$G(y_0 + \eta) = G(y_0) + G'(y_0)\,\eta + o_2(\eta)$$

in $H(x_0 + \xi) = G\big(F(x_0 + \xi)\big)$ ein. Da $F'(x_0)$ und $G'(y_0)$ beschränkte lineare Operatoren sind, ergibt sich

$$\begin{aligned} H(x_0 + \xi) &= G\big(y_0 + F'(x_0)\,\xi + o_1(\xi)\big) \\ &= G(y_0) + G'(y_0)\,\big(F'(x_0)\,\xi + o_1(\xi)\big) + o_2\big(F'(x_0)\,\xi + o_1(\xi)\big) \\ &= G(y_0) + \big(G'(y_0)\,F'(x_0)\big)\,\xi + o_3(\xi), \end{aligned}$$

was zu zeigen war.

Wenn F, G, H reelle Funktionen auf der Zahlengeraden sind, dann geht diese Formel in die Kettenregel für die Differentiation zusammengesetzter Funktionen über.

4. *Sind F und G zwei stetige Abbildungen aus X in Y und beide im Punkt x_0 differenzierbar, so sind auch die Abbildungen $F + G$ und aF (a eine Zahl) in diesem Punkt differenzierbar, und es gilt*

$$(F + G)'\,(x_0) = F'(x_0) + G'(x_0), \tag{5}$$

$$(aF')\,(x_0) = aF'(x_0). \tag{6}$$

Das folgt sofort aus den Definitionen der Summe von Operatoren und des Produktes von Operatoren mit Zahlen, wenn wir dort die Voraussetzungen einsetzen:

$$\begin{aligned} (F + G)\,(x_0 + h) &= F(x_0 + h) + G(x_0 + h) \\ &= F(x_0) + G(x_0) + F'(x_0)\,h + G'(x_0)\,h + o_1(h) \end{aligned}$$

und

$$aF(x_0 + h) = aF(x_0) + aF'(x_0)\,h + o_2(h).$$

10.1.2. Das schwache Differential (Gâteauxsches Differential). Es seien X und Y wieder zwei lineare normierte Räume, und F sei eine Abbildung aus X in Y. Existiert der Grenzwert

$$DF(x, h) = \frac{d}{dt}\,F(x + th)\bigg|_{t=0} = \lim_{t\to 0} \frac{F(x + th) - F(x)}{t}$$

(bezüglich der Normkonvergenz in Y), dann wird er *schwaches Differential* (oder *Gâteauxsches Differential*) der Abbildung F im Punkt x (bei dem Zuwachs h) genannt.

[1]) Hier und im weiteren bezeichnet der Ausdruck $o(\xi)$ Größen (aus dem entsprechenden normierten Raum, oder Zahlen), die der Bedingung $\frac{\|o(\xi)\|}{\|\xi\|} \to 0$ für $\|\xi\| \to 0$ genügen.

Das schwache Differential $DF(x, h)$ braucht nicht linear in h zu sein. Ist es jedoch linear, d. h., gilt $DF(x, h) = F_c'(x)\, h$, wobei $F_c'(x)$ ein beschränkter linearer Operator ist, dann heißt F *schwach differenzierbar* und $F_c'(x)$ die *schwache Ableitung* (oder *Gâteauxsche Ableitung*) von F an der Stelle x.

Wir weisen darauf hin, daß die Aussage über die Differentiation zusammengesetzter Abbildungen, wie sie im vorigen Abschnitt für die starke Ableitung formuliert wurde, im allgemeinen für schwache Ableitungen nicht gilt. (Man gebe ein Beispiel dafür an!)

10.1.3. Eine Abschätzung. Es sei O eine offene Menge in X und

$$[x_0, x] = \{(x_0 + t\Delta x) : \Delta x = (x - x_0),\ 0 \leqq t \leqq 1\}$$

ein Intervall, das ganz in O enthalten ist. F sei eine Abbildung, die O in Y abbildet und in jedem Punkt des Intervalls $[x_0, x]$ eine schwache Ableitung F_c' besitzt. Untersuchen wir für ein beliebiges Funktional $\varphi \in Y^*$ die Zahlenfunktion

$$f(t) = \varphi\big(F(x_0 + t\,\Delta x)\big), \qquad 0 \leqq t \leqq 1,$$

so zeigt sich, daß diese Funktion nach t im üblichen Sinn differenzierbar ist. Dazu vertauscht man in

$$\lim_{\Delta t \to 0} \frac{f(t + \Delta t) - f(t)}{\Delta t} = \lim_{\Delta t \to 0} \varphi\left(\frac{F(x_0 + t\,\Delta x + \Delta t\,\Delta x) - F(x_0 + t\,\Delta x)}{\Delta t}\right)$$

den Grenzübergang $\Delta t \to 0$ mit dem stetigen Funktional und erhält

$$f'(t) = \varphi\big(F_c'(x_0 + t\,\Delta x)\,\Delta x\big).$$

Nach dem Mittelwertsatz der Differentialrechnung ergibt sich daraus

$$f(1) - f(0) = f'(\theta),$$

d. h.

$$\varphi\big(F(x) - F(x_0)\big) = \varphi\big(F_c'(x_0 + \theta\,\Delta x)\,\Delta x\big) \tag{7}$$

mit einem geeigneten θ $(0 \leqq \theta \leqq 1)$, das natürlich auch vom Funktional φ abhängt. Für ein beliebiges Funktional $\varphi \in Y^*$ ist daher die Abschätzung

$$|\varphi\big(F(x) - F(x_0)\big)| \leqq \|\varphi\| \cdot \sup_{0 \leqq \theta \leqq 1} \|F_c'(x_0 + \theta\,\Delta x)\| \cdot \|\Delta x\| \tag{8}$$

richtig, die für das spezielle Funktional $\tilde{\varphi}$[1]) mit der Eigenschaft

$$\tilde{\varphi}\big(F(x) - F(x_0)\big) = \|\tilde{\varphi}\| \cdot \|F(x) - F(x_0)\|$$

---

[1]) Ein solches Funktional existiert nach dem Satz von Hahn-Banach.

die Ungleichung

$$\|F(x) - F(x_0)\| \leqq \sup_{0 \leqq \theta \leqq 1} \|F_c'(x_0 + \theta\, \Delta x)\| \cdot \|\Delta x\| \qquad (\Delta x = x - x_0) \tag{9}$$

liefert. Diese Ungleichung stellt das Analogon zur Abschätzung der Differenz der Werte einer differenzierbaren Zahlenfunktion dar. Die Anwendung der Abschätzung (9) auf die Abbildung

$$x \to F(x) - F_c'(x_0)\, \Delta x$$

ergibt die Ungleichung

$$\|F(x) - F(x_0) - F_c'(x_0)\, \Delta x\|$$
$$\leqq \sup_{0 \leqq \theta \leqq 1} \|F_c'(x_0 + \theta\, \Delta x) - F_c'(x_0)\| \cdot \|\Delta x\|. \tag{10}$$

10.1.4. Der Zusammenhang zwischen schwacher und starker Differenzierbarkeit. Schwache und starke Differenzierbarkeit sind verschiedene Eigenschaften sogar schon im Fall des endlichdimensionalen Raumes. So folgt für die reelle Funktion

$$f(x) = f(x_1, \ldots, x_n)$$

mehrerer unabhängiger Veränderlicher ($n \geqq 2$) aus der Existenz der Ableitung

$$\frac{d}{dt} f(x + th)$$

für ein beliebiges festes $h = (h_1, \ldots, h_n)$ nicht die (vollständige) Differenzierbarkeit dieser Funktion, d. h. die Möglichkeit, $f(x + h) - f(x)$ als Summe aus einem in h linearen Term (vollständiges Differential) und einem Rest darzustellen, der mit höherer als erster Ordnung bezüglich h verschwindet.

Ein einfaches Beispiel hierfür liefert die Funktion

$$f(x_1, x_2) = \begin{cases} \dfrac{x_1^3 x_2}{x_1^4 + x_2^2} & \text{für} \quad (x_1, x_2) \neq (0, 0), \\ 0 & \text{für} \quad (x_1, x_2) = (0, 0). \end{cases} \tag{11}$$

Sie ist in der ganzen Ebene stetig und besitzt im Punkt $(0, 0)$ für $h = (h_1, h_2)$ das schwache Differential

$$\lim_{t \to 0} \frac{f(0 + th) - f(0)}{t} = \lim_{t \to 0} \frac{t^4 h_1^3 h_2}{t^4 h_1^4 + t^2 h_2^2} = 0.$$

Damit kann das starke (vollständige) Differential im Punkt $(0, 0)$, falls es existiert, für beliebige $h = (h_1, h_2)$ höchstens den Wert 0 haben. Das würde bedeuten, daß der

Grenzwert

$$\lim_{\|h\|\to 0} \frac{f(h_1, h_2) - f(0, 0)}{\|h\|}$$

stets 0 sein müßte. Wählt man aber $h_2 = h_1^2 \neq 0$, so ist

$$\lim_{\|h\|\to 0} \frac{f(h_1, h_2) - f(0, 0)}{\|h\|} = \lim_{h_1\to 0} \frac{h_1^5}{2h_1^4 \sqrt{h_1^2 + h_1^4}} = \frac{1}{2} \neq 0,$$

d. h., im Punkt $(0, 0)$ besitzt diese Funktion $f(x_1, x_2)$ kein starkes (vollständiges) Differential, obwohl sie für jedes h ein schwaches Differential besitzt.

Umgekehrt zieht jedoch die Existenz des starken Differentials die des schwachen Differentials und der schwachen Ableitung nach sich, wobei schwache und starke Ableitung übereinstimmen. Das ergibt sich unmittelbar aus den Beziehungen

$$F(x + th) - F(x) = F'(x)\,(th) + o(th) = tF'(x)\,h + o(th)$$

und

$$\frac{F(x + th) - F(x)}{t} = F'(x)\,h + \frac{o(th)}{t} \to F'(x)\,h.$$

Im folgenden Satz beschreiben wir Bedingungen, unter denen sich diese Aussage umkehren läßt, d. h. Bedingungen, unter denen aus der schwachen Differenzierbarkeit einer Abbildung F ihre starke Differenzierbarkeit folgt.

Satz 1. *Wenn für eine Abbildung F in einer Umgebung $U(x_0)$ des Punktes x_0 die schwache Ableitung existiert und als Operatorfunktion auf $U(x_0)$ im Punkt x_0 stetig ist, dann existiert in x_0 auch die starke Ableitung $F'(x_0)$ und ist gleich der schwachen Ableitung.*

Beweis. $F_c'(x_0)$ sei die schwache Ableitung von F in x_0, d. h. $DF(x_0, h) = F_c'(x_0)\,h$, und h sei so gewählt, daß $(x_0 + th) \in U(x_0)$ für alle $t \in [0, 1]$ gilt. Für den Nachweis der Aussagen des Satzes genügt es nun zu zeigen, daß die Norm des Ausdrucks

$$\omega(x_0, h) = F(x_0 + h) - F(x_0) - F_c'(x_0)\,h \tag{12}$$

mit höherer Ordnung als $\|h\|$ gegen 0 strebt.

Ist y^* ein beliebiges Element aus dem zu Y adjungierten Raum Y^*, dann folgt aus (12)

$$\big(\omega(x_0, h), y^*\big) = \big(F(x_0 + h) - F(x_0), y^*\big) - \big(F_c'(x_0)\,h, y^*\big). \tag{13}$$

Da die Funktion $f(t) = \big(F(x_0 + th), y^*\big)$ differenzierbar und ihre Ableitung

$$\frac{df}{dt} = \lim_{\Delta t\to 0} \left(\frac{F(x_0 + th + \Delta th) - F(x_0 + th)}{\Delta t}, y^*\right) = \big(F_c'(x_0 + th)\,h, y^*\big)$$

ist, ergibt sich mit Hilfe des Mittelwertsatzes aus (13)

$$\big(\omega(x_0, h), y^*\big) = \big(\big[F_c'(x_0 + \tau h) - F_c'(x_0)\big]\,h, y^*\big), \tag{13'}$$

wobei τ eine von h abhängige Zahl zwischen 0 und 1 ist. Wählen wir nun $y^* \in Y^*$ bei festem h, so daß $\|y^*\| = 1$ und

$$\left|\left(\omega(x_0, h), y^*\right)\right| \geqq \frac{1}{2} \|\omega(x_0, h)\| \cdot \|y^*\| = \frac{1}{2} \|\omega(x_0, h)\|$$

ist[1]), so erhalten wir aus (13′) die Abschätzung

$$\|\omega(x_0, h)\| \leqq 2 \|F_c'(x_0 + \tau h) - F_c'(x_0)\| \cdot \|h\|.$$

Zusammen mit der Voraussetzung, daß $F_c'(x)$ als Operatorfunktion von x stetig ist, d. h.

$$\lim_{\|h\| \to 0} \|F_c'(x_0 + \tau h) - F_c'(x_0)\| = 0$$

ist, folgt nun aus der letzten Ungleichung

$$\lim_{\|h\| \to 0} \frac{\|\omega(x_0, h)\|}{\|h\|} = 0,$$

was zu zeigen war.

Im weiteren werden wir, wenn nicht ausdrücklich das Gegenteil vorausgesetzt wird, stets stark differenzierbare Abbildungen betrachten, die natürlich dann auch schwach differenzierbar sind.

10.1.5. Differenzierbare Funktionale. In 10.1.1. hatten wir das starke Differential für eine Abbildung F definiert, die aus einem normierten Raum X in einen anderen normierten Raum abbildet. Die Ableitung $F'(x)$ wurde dabei als linearer beschränkter Operator von X in Y, d. h. als Element des Raumes $\mathscr{L}(X, Y)$, definiert.

Ist nun Y gleich der Zahlengeraden und F ein reellwertiges Funktional auf X, so ist die Ableitung des Funktionals F in jedem Punkt x_0 ein lineares Funktional $F'(x_0)$, d. h. ein Element des Raumes X^*.

Beispiel. Wir betrachten auf dem reellen Hilbertraum H das Funktional $F(x) = \|x\|^2$. Aus der Gleichung

$$\|x + h\|^2 - \|x\|^2 = 2(x, h) + \|h\|^2$$

folgt, daß der in h lineare Teil $2(x, h)$ das starke Differential des Funktionals $F(x) = \|x\|^2$ und die Ableitung gleich dem Funktional

$$F'(x) = (2x, \cdot)$$

ist, was auch durch

$$F'(x) = 2x$$

[1]) Anm. d. Übers.: Ein solches Funktional y^* existiert nach dem Satz von HAHN-BANACH.

31*

ausgedrückt wird. Nach Satz 1 gilt

$$F'(x) = F_c'(x) = 2x.$$

Aufgabe. Man bestimme die Ableitung des Funktionals $F(x) = \|x\|$ auf einem Hilbertraum. (Lösung: $F'(x) = \frac{x}{\|x\|}$ für $x \neq 0$; im Punkt $x = 0$ existiert keine Ableitung.)

10.1.6. Abstrakte Funktionen. Es sei X jetzt der Banachraum der reellen Zahlen und Y ein beliebiger Banachraum. Eine Abbildung F aus X in Y, die einer reellen Zahl ein Element $F(x)$ des Banachraumes Y zuordnet, wird *abstrakte Funktion* genannt.

Die Ableitung $F'(x)$ einer abstrakten Funktion ist für jedes x (wo sie existiert) ein Element des Raumes Y (Tangentenvektor an die Kurve $F(x)$). Die Begriffe der starken und der schwachen Differenzierbarkeit stimmen für abstrakte Funktionen von einer reellen Variablen) überein.

10.1.7. Integrale von abstrakten Funktionen. Es sei F eine abstrakte Funktion des reellen Arguments $t \in [a, b]$ mit Werten in einem Banachraum Y. Existiert für eine Folge von immer feiner werdenden Zerlegungen des Intervalls $[a, b]$, d. h.

$$a = t_0^{(l)} < t_1^{(l)} < \cdots < t_{n_l}^{(l)} = b \qquad (l = 1, 2, \ldots),$$

$$\max_k \, (t_{k+1}^{(l)} - t_k^{(l)}) \to 0 \quad \text{für} \quad l \to \infty,$$

ein Grenzwert (in Y) der Integralsummen

$$\sum_{k=0}^{n_l-1} F(\xi_k^{(l)}) \, [t_{k+1}^{(l)} - t_k^{(l)}], \qquad \xi_k^{(l)} \in [t_k^{(l)}, t_{k+1}^{(l)}],$$

dann heißt dieser Grenzwert (ein Element aus Y) *Integral von $F(t)$ über* $[a, b]$ und wird mit

$$\int_a^b F(t) \, dt$$

bezeichnet. Analoge Überlegungen, wie sie aus der Analysis für den Fall einer reellwertigen Funktion $F(t)$ bekannt sind, zeigen: Wenn $F(t)$ auf $[a, b]$ stetig ist, dann existiert das Integral und besitzt die Eigenschaften des gewöhnlichen Riemannschen Integrals, insbesondere folgende:

1. *Wenn U eine feste lineare stetige Abbildung des Raumes Y in einem Raum Z ist, dann ist*

$$\int_a^b UF(t) \, dt = U\left(\int_a^b F(t) \, dt\right).$$

2. *Wenn $F(t)$ die Form $f(t)\,y_0$ besitzt, wobei $f(t)$ eine reellwertige Funktion und y_0 ein festes Element aus Y ist, dann ist $\int_a^b F(t)\,dt = y_0 \int_a^b f(t)\,dt$.*

3. *Es ist* $\left\| \int_a^b F(t)\,dt \right\| \leqq \int_a^b \|F(t)\|\,dt$.

Im weiteren seien nun X und Y wieder beliebige lineare normierte Räume, und $\boldsymbol{BC}(X, Y)$ sei der lineare Raum aller stetigen beschränkten[1]) Abbildungen von X in Y. Im Raum $\boldsymbol{BC}(X, Y)$ sei eine Topologie mit Hilfe der Nullumgebungen

$$U_{n,\varepsilon} = \left\{F: \sup_{\|x\|<n} \|F(x)\| < \varepsilon\right\}$$

eingeführt (auf dem Raum $\mathscr{L}(X, Y) \subset \boldsymbol{BC}(X, Y)$ aller linearen stetigen Abbildungen von X in Y stimmt diese Topologie mit der in $\mathscr{L}(X, Y)$ üblichen Normtopologie überein).

Wir betrachten eine stetige Abbildung F eines beliebigen Intervalls

$$J = [x_0, x_0 + \Delta x] = \{x: x = x_0 + t\,\Delta x, 0 \leqq t \leqq 1\}$$

des Raumes X in den Raum $\boldsymbol{BC}(X, Y)$, d. h., jedem Punkt X wird durch F eine Abbildung $F(x) \in \boldsymbol{BC}(X, Y)$ zugeordnet, die stetig vom Vektorargument $x \in J$ abhängt. Für diese Abbildung definieren wir nun durch

$$\int_{x_0}^{x_0+\Delta x} F(x)\,dx = \int_0^1 F(x_0 + t\,\Delta x)\,\Delta x\,dt \tag{14}$$

das Integral von $F(x)$ über das Intervall J. Hierbei ist $F(x_0 + t\,\Delta x)\,\Delta x$ für jedes $t \in [0, 1]$ dasjenige Element des Raumes Y, das durch die Abbildung $F(x_0 + t\,\Delta x)$ dem Element Δx zugeordnet wird. Auf Grund der Stetigkeitsvoraussetzungen existiert das Integral auf der rechten Seite von (14) und ist ein Element des Raumes Y.

Vergleichen wir nun diesen Integralbegriff für stetige Abbildungen $F(x)$ mit dem früher definierten Ableitungsbegriff, so ergibt sich formal der gleiche Zusammenhang wie in der Differential- und Integralrechnung für reelle Funktionen, nämlich:

Ist F eine Abbildung aus X in Y, die auf dem Intervall $[x_0, x_0 + \Delta x]$ eine von x stetig abhängende starke Ableitung $F'(x)$ besitzt, dann existiert das Integral $\int_{x_0}^{x_0+\Delta x} F'(x)\,dx$, und es gilt die Beziehung

$$\int_{x_0}^{x_0+\Delta x} F'(x)\,dx = F(x_0 + \Delta x) - F(x_0). \tag{15}$$

---

[1]) Die Abbildung $F: X \to Y$ heißt *beschränkt*, wenn für jede beschränkte Menge $Q \subset X$ die Menge $F(Q)$ in Y beschränkt ist. Eine nichtlineare stetige Abbildung ist nicht notwendig beschränkt.

Einerseits ist nämlich nach Definition des Integrals

$$\int_{x_0}^{x_0+\Delta x} F'(x)\,dx = \lim \sum_{k=0}^{n-1} F'(x_0 + t_k\,\Delta x)\,\Delta x\,[t_{k+1} - t_k]$$

$$= \lim \sum_{k=0}^{n-1} F'(\tilde{x}_k)\,\Delta x_k,$$

wobei $\tilde{x}_k = x_0 + t_k\,\Delta x$, $\Delta x_k = (t_{k+1} - t_k)\,\Delta x$ gesetzt wurde. Andererseits ist für eine beliebige Zerlegung des Intervalls $0 \leqq t \leqq 1$

$$F(x_0 + \Delta x) - F(x_0) = \sum_{k=0}^{n=1} F(x_0 + t_{k+1}\,\Delta x) - F(x_0 + t_k\,\Delta x)$$

$$= \sum_{k=0}^{n-1} [F(x_{k+1}) - F(x_k)].$$

Der Unterschied zwischen beiden Summen kann nach (10) durch

$$\left\| \sum_{k=0}^{n-1} [F(x_{k+1}) - F(x_k) - F'(\tilde{x}_k)\,\Delta x_k] \right\|$$

$$\leqq \|\Delta x\| \sum_{k=0}^{n-1} (t_{k+1} - t_k) \sup_{0 \leqq \theta \leqq 1} \|F'(x_k + \theta\,\Delta x_k) - F'(\tilde{x}_k)\| \tag{16}$$

abgeschätzt werden. Weil die Ableitung $F'(x)$ stetig und daher auch gleichmäßig stetig auf $[x_0, x_0 + \Delta x]$ ist, strebt die rechte Seite von (16) bei unbegrenzter Verfeinerung der Zerlegung des Intervalls $[x_0, x_0 + \Delta x]$ gegen 0. Daraus folgt die Gleichung (15).

10.1.8. Ableitungen höherer Ordnung. Es sei F eine differenzierbare Abbildung, die den linearen normierten Raum X in den linearen normierten Raum Y abbildet. Ihre Ableitung $F'(x)$ sei für jedes $x \in X$ ein Element aus $\mathscr{L}(X, Y)$, d. h., F' ist eine Abbildung des Raumes X in den Raum der linearen Operatoren $\mathscr{L}(X, Y)$. Wenn diese Abbildung differenzierbar ist, dann heißt ihre Ableitung die *zweite Ableitung* der Abbildung F und wird mit dem Symbol F'' bezeichnet, d. h., $F''(x)$ ist ein Element des Raumes $\mathscr{L}(X, \mathscr{L}(X, Y))$ der linearen Operatoren, die X in $\mathscr{L}(X, Y)$ abbilden. Im folgenden wird gezeigt, daß die Elemente dieses Raumes eine sehr bequeme und anschauliche Interpretation als sogenannte bilineare Abbildungen gestatten.

Unter einer *bilinearen Abbildung* B des Raumes X in den Raum Y versteht man eine Abbildung, die jedem geordneten Paar (x, x') von Elementen $x, x' \in X$ ein Element $y = B(x, x') \in Y$ zuordnet und dabei folgende Bedingungen erfüllt:

1. Für beliebige $x_1, x_2, x_1', x_2' \in X$ und beliebige Zahlen α, β gilt

$$B(\alpha x_1 + \beta x_2, x_1') = \alpha B(x_1, x_1') + \beta B(x_2, x_1'),$$

$$B(x_1, \alpha x_1' + \beta x_2') = \alpha B(x_1, x_1') + \beta B(x_1, x_2').$$

2. Es existiert eine positive Zahl M, so daß für alle $x, x' \in X$

$$\|B(x, x')\| \leqq M \|x\| \cdot \|x'\| \tag{17}$$

ist.

Die Bedingung 1 drückt aus, daß die Abbildung B in jedem ihrer beiden Argumente linear ist. Die Bedingung 2 ist, wie man leicht sieht, äquivalent zur Stetigkeit der Abbildung B auf der Gesamtheit ihrer Argumente.

Die kleinste Zahl M, für die die Bedingung (17) erfüllt ist, heißt *Norm* der bilinearen Abbildung B und wird mit $\|B\|$ bezeichnet.

Die linearen Operationen sind für bilineare Abbildungen in der üblichen Weise definiert und führen zu den üblichen Folgerungen. Mit diesen Operationen bilden die bilinearen Abbildungen des Raumes X in den Raum Y selbst einen linearen normierten Raum, der mit $B(X^2, Y)$ bezeichnet wird. Ist der Raum Y vollständig, dann ist auch $B(X^2, Y)$ vollständig.

Wir untersuchen nun den Zusammenhang zwischen den Räumen $\mathscr{L}(X, \mathscr{L}(X, Y))$ und $B(X^2, Y)$. Dazu bilden wir $\mathscr{L}(X, \mathscr{L}(X, Y))$ in $B(X^2, Y)$ mit der Zuordnung

$$A \xrightarrow{\varphi} B = \varphi(A), \qquad B(x, x') = (Ax)\, x' \tag{18}$$

und $B(X^2, Y)$ in $\mathscr{L}(X, \mathscr{L}(X, Y))$ mit der Zuordnung

$$B \xrightarrow{\psi} A = \psi(B), \qquad A : x \to B(x, \cdot) \in \mathscr{L}(X, Y)$$

ab. Die Abbildungen φ und ψ sind offensichtlich linear und, wie man sich leicht überlegt, invers zueinander, d. h., sie bilden $\mathscr{L}(X, \mathscr{L}(X, Y))$ und $B(X^2, Y)$ eineindeutig aufeinander ab. Für einander entsprechende Abbildungen A und B gilt

$$\|B(x, x')\| \leqq \|Ax\| \cdot \|x'\| \leqq \|A\| \cdot \|x\| \cdot \|x'\|,$$

d. h.

$$\|B\| \leqq \|A\|, \tag{19}$$

und

$$\|Ax\| = \sup_{\|x'\| \leqq 1} \|(Ax)\, x'\| = \sup_{\|x'\| \leqq 1} \|B(x, x')\| \leqq \|B\| \cdot \|x\|,$$

d. h.

$$\|A\| \leqq \|B\|, \tag{20}$$

also

$$\|A\| = \|B\|.$$

Fassen wir dieses Ergebnis mit den obigen Aussagen zusammen, so ergibt sich: Der Raum $\mathscr{L}(X, \mathscr{L}(X, Y))$ kann eineindeutig und isometrisch auf den Raum $B(X^2, Y)$ abgebildet werden.

Damit kann die zweite Ableitung $F''(x)$, nach den Betrachtungen am Anfang dieses Abschnitts ein Element aus $\mathscr{L}(X, \mathscr{L}(X, Y))$, auch als Element des Raumes $B(X^2, Y)$ aufgefaßt werden.

Zur Illustration dieser Interpretation betrachten wir ein elementares Beispiel. X und Y seien zwei euklidische Räume der Dimension m bzw. n. Jede lineare Abbildung von X in Y kann in diesem Fall als $n \cdot m$-Matrix beschrieben werden, speziell also auch die Ableitung $F'(x)$ einer Abbildung F aus X in Y. Werden in X und Y die beiden Basen

$$\{e_1, \ldots, e_m\} \quad \text{bzw.} \quad \{f_1, \ldots, f_n\}$$

ausgewählt, dann ist

$$x = e_1 x_1 + x_2 e_2 + \cdots + x_m e_m, \qquad y = y_1 f_1 + y_2 f_2 + \cdots + y_n f_n,$$

und man erhält für die Abbildung $y = F(x)$ die Form

$$\begin{aligned} y_1 &= F_1(x_1, \ldots, x_m), \\ &\cdots\cdots\cdots\cdots \\ y_n &= F_n(x_1, \ldots, x_m), \end{aligned}$$

während

$$F'(x) = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_m} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial y_n}{\partial x_1} & \frac{\partial y_n}{\partial x_2} & \cdots & \frac{\partial y_n}{\partial x_m} \end{pmatrix}$$

ist. Die zweite Ableitung $F''(x)$ wird in diesem Fall durch die Gesamtheit der $n \cdot m \cdot m$ Größen $a_{k,ij} = \dfrac{\partial^2 y_k}{\partial x_i \, \partial x_j}$ bestimmt. Diese Gesamtheit kann nun entsprechend der Formel

$$b_{k,j} = \sum_{i=1}^{m} a_{k,ij} x_i$$

als lineare Abbildung des Raumes X in den Raum $\mathscr{L}(X, Y)$ oder entsprechend der Formel

$$y_k = \sum_{i,j=1}^{m} a_{k,ij} x_i x_j'$$

als bilineare Abbildung des Raumes X in den Raum Y aufgefaßt werden.

Analog zur zweiten Ableitung einer Abbildung F von X in Y kann offensichtlich auch der Begriff der dritten, vierten, ..., allgemein der n-ten Ableitung als Ableitung der $(n-1)$-ten Ableitung der Abbildung F eingeführt werden. Die n-te Ableitung wird dann ein Element des Raumes $\mathscr{L}\big(X, \mathscr{L}\big(X, \ldots, \mathscr{L}(X, Y)\big)\big)$. Wiederholen wir die oben durchgeführten Überlegungen, so kann wieder jedem Element dieses Raumes in natürlicher Weise ein Element des Raumes $N(X^n, Y)$ der n-stelligen linearen Abbildungen von X in Y zugeordnet werden, d. h., die n-te Ableitung einer Abbildung F von X in Y kann auch als Element des Raumes $N(X^n, Y)$ aufgefaßt

werden. Dabei ist in analoger Weise unter einer *n-stelligen linearen Abbildung* $N(x', x'', \ldots, x^{(n)})$ von X in Y eine eindeutige Zuordnung der n-Tupel $(x', x'', \ldots, x^{(n)})$ von Elementen des Raumes X zu Elementen des Raumes Y zu verstehen, wobei diese Zuordnung an jeder Stelle $x^{(i)}$ (bei Fixierung der anderen Stellen) linear ist und für eine Zahl $M > 0$ die Bedingung

$$\|N(x', x'', \ldots, x^{(n)})\| \leqq M \, \|x'\| \cdot \|x''\| \cdots \|x^{(n)}\|$$

erfüllt.

10.1.9. Differentiale höherer Ordnung. Wir hatten in 10.1.1. das (starke) Differential der Abbildung F als Resultat der Anwendung des linearen Operators $F'(x)$ auf das Element $h \in X$, d. h. $dF = F'(x)\, h$, definiert. Das Differential zweiter Ordnung wird nun entsprechend durch $d^2F = F''(x)\,(h, h)$, definiert, d. h. als derjenige quadratische Ausdruck, der der Abbildung $F''(x) \in B(X^2, Y)$ entspricht. Analog wird $d^nF = F^{(n)}(x)\,(h, \ldots, h)$, d. h. das Bild $(h, h, \ldots, h) \in X \times X \times \cdots \times X = X^n$ bei der Abbildung $F^{(n)}(x)$, als Differential n-ter Ordnung bezeichnet.

10.1.10. Die Taylorsche Formel. Unter starker Differenzierbarkeit der Abbildung F verstehen wir, daß die Differenz $F(x + h) - F(x)$ als Summe aus einem in h linearen Summanden und einen Summanden dargestellt werden kann, dessen Norm schneller als $\|h\|$ gegen 0 strebt. Eine Verallgemeinerung dieser Tatsache stellt die folgende Formel dar, die eine ähnliche Struktur wie die Taylorsche Formel für reelle Funktionen besitzt.

Satz 2. *F sei eine Abbildung aus X in Y mit dem Definitionsgebiet $O \subset X$, für die die n-te Ableitung $F^{(n)}(x)$ im Punkt x existiert und auf O eine gleichmäßig stetige Funktion von x ist. Dann gilt*

$$F(x + h) - F(x) = F'(x)\, h + \frac{1}{2!} F''(x)\,(h, h) + \cdots + \frac{1}{n!} F^{(n)}(x)\,(h, \ldots, h) + \omega(x, h) \tag{21}$$

mit $\|\omega(x, h)\| = o(\|h\|^n)$.

Beweis. Wir beweisen die Aussage durch vollständige Induktion. Für $n = 1$ ist (21) trivial. Für den Fall $n - 1$ setzen wir (21) als richtig voraus. Daher gilt für die Abbildung $F'(x)$

$$F'(x + h) = F'(x) + F''(x)\, h + \frac{1}{2!} F'''(x)\,(h, h) + \cdots + \frac{1}{(n-1)!} F^{(n)}(x)\,(h, \ldots, h) + \omega_1(x, h). \tag{22}$$

wobei $\|\omega_1(x, h)\| = o(\|h\|^{n-1})$ ist. Integrieren wir beide Seiten der Gleichung (22) über das Intervall $[x, x + h]$, dann erhalten wir nach Formel (15)

$$F(x + h) - F(x) = \int_0^1 F'(x + th)\, h\, dt$$

$$= \int_0^1 \Big\{F'(x) + tF''(x)\, h + \cdots + \frac{1}{2!}\, t^2 F'''(x)\,(h, h) + \cdots$$

$$+ \frac{1}{(n-1)!}\, t^{n-1} F^{(n)}\,(h, \ldots, h)\Big\}\, h\, dt + R_n, \tag{23}$$

wobei $R_n = \int_0^1 \omega_1(x, th)\, h\, dt$ ist. Aus (23) ergibt sich

$$F(x + h) - F(x) = F'(x)\, h + \frac{1}{2!}\, F''(x)\,(h, h) + \cdots$$

$$+ \frac{1}{n!}\, F^{(n)}(x)\,(h, \ldots, h) + R_n, \tag{24}$$

was die Aussage (21) im Fall n liefert, denn es ist

$$\|R_n\| \leqq \int_0^1 \|\omega_1(x, th)\| \cdot \|h\|\, dt = o(\|h\|^n).$$

Damit ist der Satz vollständig bewiesen.

Formel (21) wird *Taylorsche Formel für Abbildungen* genannt.

10.2. Extremalaufgaben

Einer der ältesten und sehr gut ausgearbeiteten Zweige der nichtlinearen Funktionalanalysis ist die Bestimmung der Extrema von Funktionalen. Die Untersuchung solcher Probleme bildet den Inhalt der sogenannten *Variationsrechnung*. Die Mehrzahl der Methoden der Variationsrechnung benutzt die spezielle Form der Funktionale, deren Extremwerte bestimmt werden sollen. Einige allgemeine Verfahren und Resultate können jedoch auch für mehr oder weniger beliebige Funktionale formuliert werden. Wir werden im folgenden keine vollständige Darlegung der Methoden der Variationsrechnung geben, sondern uns auf eine kurze Betrachtung von Elementen einer allgemeinen Theorie beschränken, auf denen die Variationsrechnung aufbaut.

10.2.1. Eine notwendige Bedingungen für die Existenz eines Extremums. Es sei F ein reelles Funktional auf dem Banachraum X. Ist für alle x einer beliebig kleinen Umgebung von x_0 stets $F(x) - F(x_0) \geqq 0$, so sagt man: Das Funktional F *nimmt*

in x_0 ein Minimum an. Der Begriff des Maximums von F wird entsprechend definiert. Nimmt ein Funktional F in einem Punkt x_0 ein Minimum oder ein Maximum an, dann sagt man: Das Funktional F besitzt im Punkt x_0 ein *Extremum.*

Die Bestimmung der Extrema von Funktionalen ist oft die mathematische Aufgabe bei der Lösung vieler Probleme aus Physik und Technik.

Für reelle Funktionen von n reellen Variablen ist die folgende notwendige Bedingung für die Existenz eines Extremums bekannt: Wenn die Funktion f in dem Punkt $x_0 = (x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)})$ differenzierbar ist und dort ein Extremum hat, dann ist in diesem Punkt $df = 0$ oder, was gleichbedeutend ist,

$$\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2} = \cdots = \frac{\partial f}{\partial x_n} = 0.$$

Diese Bedingung kann leicht auf Funktionale übertragen werden, die auf einem beliebigen normierten Raum definiert sind.

Satz 1. *Ist F ein differenzierbares Funktional, dann ist für die Existenz eines Extremums im Punkt x_0 notwendig, daß das Differential in diesem Punkt für alle h verschwindet, d. h.*

$$F'(x_0)\,h \equiv 0.$$

Beweis. Entsprechend der Definition der Differenzierbarkeit ist

$$F(x_0 + h) - F(x_0) = F'(x_0)\,h + o(h). \tag{1}$$

Nehmen wir an, daß F in x_0 ein Extremum annimmt und $F'(x_0)\,\tilde{h} \neq 0$ für irgendein $\tilde{h}$ ist, dann hat für betragsmäßig hinreichend kleine reelle Zahlen λ der ganze Ausdruck $F'(x_0)(\lambda\tilde{h}) + o(\lambda\tilde{h})$ das gleiche Vorzeichen wie sein Hauptteil $F'(x_0)(\lambda\tilde{h})$. Wegen $F'(x_0)\,(\lambda\tilde{h}) = \lambda F'(x_0)\,\tilde{h}$ $\big(F'(x_0)$ ist ein lineares Funktional$\big)$ wechselt dieser Ausdruck aber mit λ sein Vorzeichen und daher auch die linke Seite von (1), d. h., im Punkt x_0 liegt kein Extremum vor. Das ist ein Widerspruch zur Voraussetzung.

Wir betrachten nun einige Beispiele.

1. Wird durch

$$F(x) = \int_a^b f\big(t, x(t)\big)\,dt, \tag{2}$$

wobei f eine Funktion mit stetigen partiellen Ableitungen ist, ein Funktional F auf dem Raum $C[a, b]$ der auf $[a, b]$ stetigen Funktionen definiert, dann ist dieses Funktional differenzierbar. Denn es ist

$$F(x + h) - F(x) = \int_a^b [f(t, x + h) - f(t, x)]\,dt$$

$$= \int_a^b f_x'(t, x)\,h(t)\,dt + o(h)$$

und daher

$$dF = \int_a^b f_x'(t, x(t))\, h(t)\, dt.$$

Die notwendige Bedingung für die Existenz eines Extremums, $dF \equiv 0$ für alle $h \in C[a, b]$, führt jetzt zur Forderung $f_x'(t, x) = 0$. Das ergibt sich aus Folgendem. Für jedes $x(t) \in C[a, b]$ ist die partielle Ableitung $f_x'(t, x)$ eine stetige Funktion von t. Nehmen wir an, daß $f_x'(t, x)$ in einem Punkt t_0 verschieden von Null ist, z. B. $f_x'(t_0, x(t_0)) > 0$, so gilt diese Ungleichung auch in einer kleinen Umgebung (α, β) des Punktes t_0, und wir erhalten für die Funktion

$$\tilde{h}(t) = \begin{cases} (t - \alpha)(\beta - t) & \text{für } \alpha \leqq t \leqq \beta, \\ 0 & \text{sonst} \end{cases}$$

aus $C[a, b]$ die Ungleichung

$$\int_a^b f_x'(t, x)\, h(t)\, dt > 0.$$

Das widerspricht der Voraussetzung $dF \equiv 0$ für alle $h \in C[a, b]$.

Die Forderung $f_x'(t, x) = 0$ definiert im allgemeinen eine Kurve, auf der das Funktional (2) ein Extremum annehmen kann.

2. Im gleichen Raum $C[a, b]$ betrachten wir nun das durch die Formel

$$F(x) = \int_a^b \int_a^b K(\xi_1, \xi_2)\, x(\xi_1)\, x(\xi_2)\, d\xi_1\, d\xi_2, \tag{3}$$

definierte Funktional F. Dabei ist $K(\xi_1, \xi_2)$ eine stetige Funktion, für die $K(\xi_1, \xi_2) = K(\xi_2, \xi_1)$ gilt. Man rechnet leicht nach, daß das Differential dieses Funktionals gleich

$$dF = 2 \int_a^b \int_a^b K(\xi_1, \xi_2)\, x(\xi_1)\, h(\xi_2)\, d\xi_1\, d\xi_2$$

ist. Die notwendige Bedingung für die Existenz eines Extremums, $dF \equiv 0$ für alle $h \in C[a, b]$, führt jetzt, wie man analog zu Beispiel 1 beweist, zur Forderung

$$\int_a^b K(\xi_1, \xi_2)\, x(\xi_1)\, d\xi_1 = 0 \quad \text{für alle} \quad \xi_2 \in [a, b].$$

Eine der Lösungen dieser Gleichung ist die Funktion $x \equiv 0$. Die Frage, ob in diesem Punkt ein Extremum vorliegt oder ob es noch andere Punkte gibt, in denen Extrema des Funktionals möglich sind, kann nur bei Kenntnis der konkreten Form des Kerns durch zusätzliche Untersuchungen beantwortet werden.

3. Wir betrachten nun das durch die Formel

$$F(x) = \int_a^b f\big(t, x(t), x'(t)\big)\,dt \tag{4}$$

definierte Funktional F auf dem Raum $C^1[a, b]$ der auf $[a, b]$ stetig differenzierbaren Funktionen. Dabei ist $x'(t) = \dfrac{dx(t)}{dt}$, und $f(t, x, x')$ ist eine nach allen Argumenten zweimal stetig differenzierbare Funktion. Dieses Funktional spielt eine grundlegende Rolle in vielen Fragen der Variationsrechnung.

Mit Hilfe der Taylorschen Formel erhalten wir

$$\begin{aligned} F(x + h) - F(x) &= \int_a^b [f(t, x + h, x' + h') - f(t, x, x')]\,dt \\ &= \int_a^b (f_x h + f_{x'} h')\,dt + o(\|h\|), \end{aligned}$$

wobei $\|h\|$ die Norm der Funktion h als Element des Raumes $C^1[a, b]$ ist. Damit hat die notwendige Bedingung für die Existenz eines Extremums des Funktionals (4) die Gestalt

$$dF = \int_a^b (f_x h + f_{x'} h')\,dt = 0 \tag{5}$$

Diese integrale Formulierung der Bedingung ist zur Bestimmung der Funktion x, für die das Funktional das Extremum annimmt, wenig geeignet. Wir formen sie daher in eine für die Bestimmung der Lösung günstigere Form um. Integrieren wir in (5) den Term $f_{x'} h'$ partiell,[1] so ergibt sich

$$\int_a^b f_{x'} h'\,dt = f_{x'} h\Big|_a^b - \int_a^b h \frac{d}{dt} f_{x'}\,dt$$

und, wenn wir diese Beziehung in (5) einsetzen,

$$dF = \int_a^b \left(f_x - \frac{d}{dt} f_{x'}\right) h\,dt + f_{x'} h\,\Bigg|_a^b = 0. \tag{6}$$

Da diese Gleichung für alle $h \in C^1[a, b]$ erfüllt sein muß, insbesondere auch für Funktionen h mit der Eigenschaft $h(a) = h(b) = 0$, muß andererseits

$$\int_a^b \left(f_{x'} - \frac{d}{dt} f_{x'}\right) h\,dt = 0$$

für alle $h \in C^1[a, b]$ mit $h(a) = h(b) = 0$ gelten.

[1] Diese Operation erfordert eine zusätzliche Begründung, da die Existenz der im Ausdruck $\dfrac{d}{dt} f_{x'}$ auftretenden zweiten Ableitung x'' nicht vorausgesetzt ist. Man findet eine solche Begründung in jedem Buch über Variationsrechnung.

Durch analoge Überlegungen wie im Beispiel 1 leiten wir daraus die Forderung ab, daß, um die Bedingung (6) für alle $h \in C^1[a, b]$ zu erfüllen, mindestens

$$f_x - \frac{d}{dt} f_{x'} = 0 \tag{7}$$

und (wie aus (6) sofort folgt) zusätzlich

$$f_{x'} \cdot h \,\Big|_a^b = 0 \tag{8}$$

sein muß. Betrachtet man die letzte Bedingung für zwei spezielle Funktionenklassen $\tilde{h}$ bzw. $\tilde{\tilde{h}}$ aus $C^1[a, b]$ mit $\tilde{h}(a) = 0$, $\tilde{h}(b) \neq 0$ bzw. $\tilde{\tilde{h}}(a) \neq 0$, $\tilde{\tilde{h}}(b) = 0$, so ergeben sich

$$f_{x'} \,\Big|_{t=b} = 0 \tag{9}$$

und

$$f_{x'} \,\Big|_{t=a} = 0 \tag{10}$$

als Zusatzbedingungen. Aus der notwendigen Bedingung (6) für die Existenz eines Extremums des Funktionals (4) folgt also, daß die Funktion $x \in C^1[a, b]$, für die das Funktional ein Extremum annimmt, notwendig die Differentialgleichung (7) und die Randbedingungen (9) und (10) in den Endpunkten des Intervalls $[a, b]$ erfüllen muß. Da die allgemeine Lösung einer Differentialgleichung zweiter Ordnung zwei freie Konstanten enthält, stehen mit den Randbedingungen (9), (10) hier gerade soviel Bedingungen bereit, wie zur Bestimmung dieser Konstanten notwendig sind.

10.2.2. Das zweite Differential. Hinreichende Bedingungen für die Existenz eines Extremums. Ausgangspunkt unserer Betrachtungen ist der einfache Fall der Bestimmung des Extremums einer Funktion von n Veränderlichen. Ist für die Funktion $f(x_1, \ldots, x_n)$ im Punkt $(x_1^{(0)}, \ldots, x_n^{(0)})$ die Bedingung $df = 0$ erfüllt, dann kann an dieser Stelle von f ein Extremum angenommen werden. Zur Entscheidung der Frage, ob an dieser Stelle wirklich ein Extremum der Funktion f vorliegt oder nicht, kann das zweite Differential von f herangezogen werden. Es ist bekannt:

1. *Wenn die Funktion $f(x_1, \ldots, x_n)$ im Punkt $(x_1^{(0)}, \ldots, x_n^{(0)})$ ein Minimum hat, dann ist in diesem Punkt $d^2f \geqq 0$.* Analog: *Wenn im Punkt $(x_1^{(0)}, \ldots, x_n^{(0)})$ ein Maximum vorliegt, dann ist dort $d^2f \leqq 0$.*

2. *Wenn im Punkt $(x_1^{(0)}, \ldots, x_n^{(0)})$ die Bedingungen*

$$df = 0 \quad \textit{und} \quad d^2f = \sum_{i,k=1}^{n} \frac{\partial^2 f}{\partial x_i \, \partial x_k} \, dx_i \, dx_k > 0 \tag{11}$$

(wenn nicht alle $dx_i = 0$) *erfüllt sind, dann nimmt $f(x)$ in diesem Punkt ein Minimum an* (analog: ein Maximum, wenn $d^2f < 0$ ist).

Kürzer gesagt, die Nichtnegativität des zweiten Differentials ist notwendig, seine positive Definitheit ist hinreichend für die Existenz eines Minimums.

Satz 2. *F sei ein reelles Funktional auf dem Banachraum X, das in einer Umgebung des Punktes x_0 eine stetige zweite Ableitung besitzt. Wenn dieses Funktional im Punkt x_0 ein Minimum annimmt, dann ist $d^2F(x_0) \geqq 0$.*[1])

Beweis. Nach der Taylorschen Formel ist

$$F(x_0 + h) - F(x_0) = F'(x_0)\, h + \frac{1}{2} F''(x_0)\,(h, h) + o(\|h\|^2).$$

Wenn das Funktional F im Punkt x_0 ein Minimum hat, dann ist $F'(x_0) = 0$, und es gilt

$$F(x_0 + h) - F(x_0) = \frac{1}{2} F''(x_0)\,(h, h) + o(\|h\|^2). \tag{12}$$

Nehmen wir nun an, daß für ein h die Ungleichung

$$F''(x_0)\,(h, h) < 0 \tag{13}$$

erfüllt ist, dann ist sie es auf Grund von $F''(x_0)\,(\varepsilon h, \varepsilon h) = \varepsilon^2 F''(x_0)\,(h, h)$ auch für Elemente h mit beliebig kleiner Norm. Da aber für hinreichend kleine $\|h\|$ das Vorzeichen des ganzen Ausdrucks (12) mit dem seines Hauptteils $\frac{1}{2} F''(x_0)\,(h, h)$ übereinstimmt, erhalten wir

$$F(x_0 + h) - F(x_0) = \frac{1}{2} F''(x_0)\,(h, h) + o(\|h\|^2) < 0,$$

d. h., unter der Annahme (13) kann in x_0 kein Minimum vorliegen. Dieser Widerspruch zur Voraussetzung zeigt die Richtigkeit des Satzes.

Ein analoger Satz kann für das Maximum formuliert werden.

Der eben bewiesene Satz ist eine direkte Verallgemeinerung des entsprechenden Satzes für reelle Funktionen endlich vieler Veränderlicher. Anders verhält es sich bei der hinreichenden Bedingung. Es zeigt sich, daß die der Bedingung (11) entsprechende Bedingung $F''(x_0)\,(h, h) > 0$ für Funktionale auf einem unendlichdimensionalen Banachraum nicht hinreichend ist. Dazu betrachten wir folgendes Beispiel. Im Hilbertraum l_2 sei das Funktional

$$F(x) = \sum_{n=1}^{\infty} \frac{x_n^2}{n^3} - \sum_{n=1}^{\infty} x_n^4$$

gegeben. Im Punkt 0 ist das erste Differential dieses Funktionals gleich 0, das zweite Differential gleich der Reihe $2 \sum_{n=1}^{\infty} \frac{h_n^2}{n^3}$, d. h. selbst ein positiv definites Funktional. Dennoch nimmt das Funktional F im Punkt 0 kein Minimum an, da aus $F(0) = 0$ und $F\Big(\underbrace{0, \ldots, 0}_{n-1}, \frac{1}{n}, 0, \ldots\Big) = \frac{1}{n^5} - \frac{1}{n^4} < 0$ folgt, daß es in jeder Umgebung des Punktes 0 noch Punkte gibt, in denen $F(x) < F(0)$ ist.

[1]) Diese Ungleichung besagt, daß $F''(x_0)\,(h, h) \geqq 0$ für alle h ist.

Zur Formulierung einer hinreichenden Bedingung für die Existenz eines Extremums führen wir folgenden Begriff ein. Ein quadratisches Funktional B heißt *stark positiv*, wenn es eine Konstante $c > 0$ gibt, so daß $B(x, x) \geqq c \, \|x\|^2$ für alle x ist.

Satz 3. *Wenn das Funktional F auf dem Banachraum X die Bedingungen*

1. $dF(x_0) = 0$,

2. $d^2F(x_0)$ *ist ein stark positives quadratisches Funktional,*

erfüllt, dann nimmt das Funktional F im Punkt x_0 ein Minimum an.

Beweis. Es sei $\varepsilon > 0$ eine so kleine Zahl, daß für $\|h\| < \varepsilon$ die Größe $o(\|h\|^2)$ in Gleichung (12) die Bedingung

$$|o(\|h\|^2)| < \frac{c}{4} \|h\|^2$$

erfüllt. Dann ist

$$F(x_0 + h) - F(x_0) = \frac{1}{2} F''(x_0) (h, h) + o(\|h\|^2) > \frac{c}{4} \|h\|^2 > 0 \quad \text{für } \|h\| < \varepsilon,$$

was zu zeigen war.

Im endlichdimensionalen Raum ist die starke Positivität eines quadratischen Funktionals äquivalent zu seiner positiven Definitheit, daher ist (bei Verschwinden des ersten Differentials) die positive Definitheit des zweiten Differentials hinreichend für die Existenz eines Minimums der Funktion. Im unendlichdimensionalen Fall ist (wie das oben angegebene Beispiel zeigt) die starke Positivität eine stärkere Bedingung als die positive Definitheit.

Die Bedingung der starken Positivität des zweiten Differentials, die ein Minimum garantiert, ist einerseits sehr bequem, da sie auf ein beliebiges zweimal differenzierbares Funktional (unabhängig von seiner konkreten Form) in einem beliebigen Banachraum angewandt werden kann. Andererseits erweist sich diese Bedingung in praktisch wichtigen Fällen als zu grob und zu schwierig nachweisbar. In der Variationsrechnung werden daher engere hinreichende Bedingungen für die Existenz eines Extremums aufgestellt, die von der konkreten Form der in den Variationsaufgaben benutzten Funktionale ausgehen. Eine Darstellung dieser Fragen ist jedoch nicht Anliegen dieses Buches.

10.3. Das Newtonsche Verfahren

Eine der bekanntesten Methoden zur Lösung von Gleichungen der Form

$$f(x) = 0 \tag{1}$$

(f eine reellwertige Funktion einer reellen Variablen $x \in [a, b]$) ist das sogenannte *Newtonsche Verfahren* oder *Tangentenverfahren*. Es besteht darin, daß nach der

Rekursionsformel

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \tag{2}$$

(ausgehend von einem beliebigen Punkt $x_0 \in [a, b]$) Näherungen für die Lösung berechnet werden. Abb. 24 gibt dazu eine geometrische Veranschaulichung. Besitzt die Funktion f auf dem Intervall $[a, b]$ nur eine einzige Nullstelle x^*, eine nirgends verschwindende erste und eine beschränkte zweite Ableitung, so kann man zeigen, daß es eine Umgebung $U(x^*)$ des Punktes x^* mit folgender Eigenschaft gibt: Wird ein beliebiger Punkt aus $U(x^*)$ als nullte Näherung x_0 in die Formel (2) eingesetzt, dann konvergiert die nach (2) berechnete Näherungsfolge $\{x_n\}$ gegen die Nullstelle x^*.

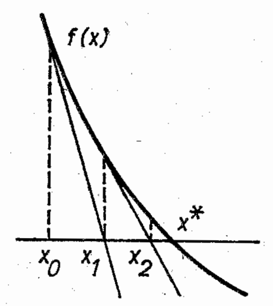


Abb. 24

Das Newtonsche Verfahren kann auf Operatorgleichungen übertragen werden. Wir führen dies hier für Gleichungen in einem Banachraum durch. Dazu betrachten wir die Gleichung

$$F(x) = 0, \tag{3}$$

wobei F eine Abbildung des Banachraumes X in den Banachraum Y ist. Die Abbildung F sei stark differenzierbar in einer Kugel $B(x_0, r)$ mit dem Radius r um den Mittelpunkt x_0, den wir als nullte Näherung der gesuchten Lösung verwenden. Ersetzen wir, wie im eindimensionalen Fall, den Ausdruck $F(x) - F(x_0)$ durch seinen linearen Hauptteil, d. h. durch das Element $F'(x_0)(x - x_0)$, dann erhalten wir aus (3) die lineare Gleichung

$$F'(x_0)(x_0 - x) = F(x_0),$$

deren Lösung

$$x_1 = x_0 - [F'(x_0)]^{-1}\big(F(x_0)\big)$$

wir als nächste Näherung der exakten Lösung von (3) ansehen (dabei wird natürlich die Existenz des Operators $[F'(x_0)]^{-1}$ vorausgesetzt). Wiederholen wir dieselben Überlegungen, so erhalten wir eine Folge

$$x_{n+1} = x_n - [F'(x_n)]^{-1}\big(F(x_n)\big) \tag{4}$$

von Näherungslösungen der Gleichung (3). Da im unendlichdimensionalen Fall die Bestimmung des Operators $[F'(x_n)]^{-1}$ eine sehr schwierige Aufgabe sein kann, modifizieren wir diese Methode etwas, indem wir an Stelle von (4) die Formel

$$x_{n+1} = x_n - [F'(x_0)]^{-1} \big(F(x_n)\big) \tag{5}$$

benutzen. Dabei wird also in jedem Schritt der Umkehroperator $[F'(x_0)]^{-1}$ für ein und dasselbe Argument x_0 benutzt. Diese Modifikation gegenüber (3) vermindert zwar etwas die Konvergenzgeschwindigkeit des Verfahrens, erweist sich aber vom rechentechnischen Standpunkt aus oft als zweckmäßiger. Man bezeichnet das Näherungsverfahren, das die Formel (5) benutzt, als *modifiziertes Newtonsches Verfahren* (vgl. [29], [30]).

Wir kommen nun zur präzisen Formulierung des modifizierten Newtonschen Verfahrens.

Satz 1. *Ist F eine Abbildung, die in einer Kugel $B(x_0, r)$ mit dem Mittelpunkt x_0 und dem Radius r stark differenzierbar ist und deren Ableitung in dieser Kugel einer Lipschitzbedingung*

$$\|F'(x_1) - F'(x_2)\| \leqq L\|x_1 - x_2\|$$

genügt, existiert ferner der Umkehroperator $[F'(x_0)]^{-1}$ und ist

$$M = \|[F'(x_0)]^{-1}\|, \qquad k = \|[F'(x_0)]^{-1} \big(F(x_0)\big)\|, \qquad h = MkL < \frac{1}{4},$$

dann besitzt die Gleichung $F(x) = 0$ in der Kugel $\|x - x_0\| < kt_0$ (t_0 kleinste Wurzel der Gleichung $ht^2 - t + 1 = 0$) eine eindeutig bestimmte Lösung x^*. Diese Lösung erhält man als Grenzwert der durch die Rekursionsformel (5) *definierten Folge* $\{x_n\}$.

Beweis. Wir betrachten auf dem Hilbertraum X die Abbildung

$$Ax = x - [F'(x_0)]^{-1} \big(F(x)\big).$$

Ihre starke Ableitung im Punkt x_0 ist gleich 0. Die Kugel $\|x - x_0\| \leqq kt_0$ wird bei dieser Abbildung auf sich selbst abgebildet. Denn es ist

$$\begin{aligned} Ax - x_0 &= x - x_0 - [F'(x_0)]^{-1} \big(F(x)\big) \\ &= [F'(x_0)]^{-1} \{F'(x_0)(x - x_0) - F(x) + F(x_0)\} - [F'(x_0)]^{-1} \big(F(x_0)\big). \end{aligned}$$

Daraus folgt

$$\begin{aligned} \|Ax - x_0\| \leqq{} & \|[F'(x_0)]^{-1}\| \cdot \|F'(x_0)(x - x_0) - F(x) + F(x_0)\| \\ & + \|[F'(x_0)]^{-1} F(x_0)\|, \end{aligned}$$

d. h.

$$\|Ax - x_0\| \leqq M \|F'(x_0)(x - x_0) - F(x) + F(x_0)\| + k. \tag{6}$$

Da andererseits die Abbildung $\Phi(x) = F(x) - F(x_0) - F'(x_0)(x - x_0)$ differenzierbar ist und ihre Ableitung $\Phi'(x) = F'(x) - F'(x_0)$ für $\|x - x_0\| \leqq Lt_0k$ der Abschätzung

$$\|\Phi'(x)\| = \|F'(x) - F'(x_0)\| \leqq L\|x - x_0\| \leqq Lt_0k$$

genügt, folgt aus dem Mittelwertsatz (vgl. (9) in 10.1.)

$$\|\Phi(x)\| = \|\Phi(x) - \Phi(x_0)\| \leqq Lt_0k \|x - x_0\| \leqq Lt_0^2k^2. \tag{7}$$

Damit ergibt sich für $\|x - x_0\| \leqq t_0k$ aus (6)

$$\|Ax - x_0\| \leqq MLt_0^2k^2 + k = k(MLt_0^2k + 1) = k(ht_0^2 + 1) = kt_0,$$

d. h., A bildet die Kugel $\|x - x_0\| \leqq k_0$ auf sich selbst ab. Wir zeigen nun, daß A eine kontrahierende Abbildung auf dieser Kugel ist. Für $\|x - x_0\| \leqq kt_0$ ist

$$A'(x) = I - [F'(x_0)]^{-1}(F'(x)) = [F'(x_0)]^{-1}\big(F'(x_0) - F'(x)\big)$$

und deshalb

$$\|A'(x)\| \leqq M\|F'(x_0) - F'(x)\| \leqq ML\|x - x_0\| \leqq MLkt_0.$$

Da t_0 nach Voraussetzung die kleinste Wurzel der Gleichung $ht^2 - t + 1 = 0$, d. h. $t_0 = \dfrac{1 - \sqrt{1 - 4h}}{2h}$ ist, folgt

$$\|A'(x)\| \leqq MLkt_0 = ht_0 = h\,\frac{1 - \sqrt{1 - 4h}}{2h} = \frac{1 - \sqrt{1 - 4h}}{2} = q < \frac{1}{2}. \tag{8}$$

Daher ist nach dem Mittelwertsatz

$$\|Ax_1 - Ax_2\| < \frac{1}{2}\|x_1 - x_2\|,$$

d. h., A ist auf der Kugel $\|x - x_0\| \leqq kt_0$ eine kontrahierende Abbildung. Folglich besitzt A dort einen eindeutig bestimmten Fixpunkt x^*. Für diesen Punkt gilt

$$x^* = x^* - [F'(x_0)]^{-1}\big(F(x^*)\big), \quad \text{d. h. } F(x^*) = 0.$$

Aus dem Satz über die kontrahierenden Abbildungen (vgl. Satz 1 aus 2.4.) folgt, daß die nach der Rekursionsformel

$$Ax_n = x_n - [F'(x_0)]^{-1}\big(F(x_0)\big) = x_{n+1}$$

gebildete Folge $\{x_n\}$ gegen x^* konvergiert, was zu zeigen war.

Aus der Ungleichung (8) ergibt sich sofort folgende Abschätzung für die Konvergenzgeschwindigkeit des modifizierten Newtonschen Verfahrens:

$$\|x_n - x^*\| \leqq \frac{q^n}{1-q}\,\|[F'(x_0)]^{-1}\big(F(x_0)\big)\|, \tag{9}$$

d. h., der Fehler der Näherung x_n, die nach dem modifizierten Newtonschen Verfahren (5) berechnet wird, nimmt mit geometrischer Progression ab. Der Fehler der Näherung $\tilde{x}_n$, die nach dem Newtonschen Verfahren (4) berechnet wird, kann durch

$$\|\tilde{x}_n - x^*\| \leqq \frac{1}{2^{n-1}}(2h)^{2^n-1}k$$

abgeschätzt werden, d. h., dieser Fehler nimmt stärker als mit geometrischer Progression ab.

Beispiel. Wir betrachten die nichtlineare Integralgleichung

$$x(s) = \int_a^b K(s, t, x(t))\, dt, \tag{10}$$

wobei $K(s, t, u)$ eine stetige und nach allen ihren Argumenten stetig differenzierbare Funktion ist. Wenn wir durch die Formel

$$y(s) = x(s) - \int_a^b K(s, t, x(t))\, dt$$

die Abbildung $y = F(x)$ definieren, können wir die Gleichung (10) in der Form $F(x) = 0$ schreiben.

Es sei x_0 die nullte Näherung für die Lösung dieser Gleichung. Dann wird die erste Verbesserung $\Delta x(s) = x_1 - x_0$ durch die Gleichung

$$F'(x_0)\, \Delta x = -F(x_0) \tag{11}$$

bestimmt.

Wenn die Funktion $K(s, t, u)$ und der Funktionenraum, in dem die Gleichung (10) betrachtet wird, so beschaffen sind, daß die Ableitung $F'(x)$ der Abbildung F durch „Differentiation unter dem Integralzeichen" berechnet werden kann, d. h., wenn

$$z = F'(x_0)\, x$$

durch

$$z(s) = x(s) - \int_a^b K_u'(s, t, x_0(t))\, x(t)\, dt$$

beschrieben wird, dann nimmt Gleichung (11) die Form

$$\Delta x(s) = \int_a^b K_u'(s, t, x_0(t))\, \Delta x(t)\, dt + \varphi_0(s) \tag{12}$$

an, wobei $\varphi_0(s) = \int_a^b K(s, t, x_0(t))\, dt - x_0(s)$ ist. Die weiteren Verbesserungen der Näherung findet man analog.

Auf diese Weise ist die Bestimmung jeder weiteren Näherung auf die Lösung einer linearen Integralgleichung zurückgeführt. Wenn man das modifizierte Newtonsche Verfahren anwendet, dann ist bei jedem Schritt eine lineare Integralgleichung mit ein und demselben Kern zu lösen.

Eine detaillierte Darstellung des Newtonschen Verfahrens und der damit zusammenhängenden Fragen findet sich im Buch [30] und im Artikel [29] von L. V. Kantorovič, der auch die grundlegenden Resultate des Newtonschen Verfahrens für Operatorgleichungen gefunden hat.

Anhang

Banachsche Algebren

V. M. Tichomirov

Im dritten Kapitel wurden lineare Räume untersucht. Dabei wurde eine wichtige Klasse linearer Räume, die der Banachräume, eingeführt. In diesem Anhang werden Banachsche Algebren (normierte Algebren) untersucht. Hierbei handelt es sich um Banachräume, bei denen eine Multiplikation der Elemente erklärt ist. Zusammen mit der linearen und der metrischen Struktur werden die Banachschen Algebren durch eine derartige Multiplikation mit einer Reihe bemerkenswerter Eigenschaften ausgestattet.

A.1. Definition und Beispiele Banachscher Algebren

A.1.1. Banachsche Algebren. Isomorphie Banachscher Algebren. Wir erinnern daran, daß ein linearer Raum eine nichtleere Menge von Elementen ist, in der zwei Operationen erklärt sind, eine Addition und eine Multiplikation mit Zahlen (vgl. 3.1.).

Definition 1. Einen linearen Raum X nennt man eine *Algebra*, wenn in ihm eine zusätzliche Operation, eine Multiplikation, erklärt ist, die folgenden Axiomen genügt:

1. $(xy)\,z = x(yz)$.

2. $x(y + z) = xy + xz$; $(y + z)\,x = yx + zx$.

3. $\alpha(xy) = (\alpha x)\,y = x(\alpha y)$.

4. Falls ein Element $e \in X$ existiert, so daß $ex = xe = x$ für alle $x \in X$ gilt, nennt man e *Einselement*. Die Algebra X heißt *Algebra mit Einselement*.[1])

5. Falls die Multiplikation kommutativ ist, d. h. $xy = xy$ gilt, nennt man X eine *kommutative Algebra*.

Wir werden uns hauptsächlich mit kommutativen Algebren mit Einselement beschäftigen. Hierbei betrachten wir die Algebra über dem Körper $\boldsymbol{C}$ der komplexen Zahlen.

In 3.3.1. wurde der Begriff des normierten Raumes eingeführt. Seine Norm wird mit $\|x\|$ bezeichnet.

[1]) Das Einselement einer Algebra ist stets eindeutig bestimmt. Ist nämlich e' ein Element, das ebenfalls die Voraussetzungen von Axiom 4 erfüllt, so gilt $ee' = e = e'$.

Definition 2. Ein normierter Raum X heißt *normierte Algebra*, falls er eine Algebra mit Einselement ist und die beiden Axiome

6. $\|e\| = 1$,

7. $\|xy\| \leqq \|x\| \cdot \|y\|$

erfüllt. Falls die normierte Algebra X vollständig ist (also ein Banachraum ist), heißt sie *Banachsche Algebra.*

Eine Abbildung $F: X \to Y$ nennt man einen *Homomorphismus* der Algebra X in die Algebra Y, falls

$$F(x + y) = Fx + Fy, \tag{1}$$

$$F(\alpha x) = \alpha Fx, \tag{2}$$

$$F(xy) = Fx \cdot Fy \tag{3}$$

gilt. Zwei Algebren X und Y heißen *isomorph,* falls es eine eineindeutige Abbildung F von X auf Y mit den Eigenschaften (1) bis (3) gibt.

Zwei normierte Räume X und Y heißen *isometrisch,* falls es eine eineindeutige Abbildung $F: X \leftrightarrow Y$ mit den Eigenschaften (1), (2) und

$$\|Fx\|_Y = \|x\|_X$$

gibt.

Definition 3. Zwei Banachsche Algebren heißen *isometrisch-isomorph,* falls ein Isomorphismus $F: X \leftrightarrow Y$ existiert, der eine isometrische Abbildung von X auf Y (betrachtet als normierte Räume) vermittelt.

A.1.2. Beispiele Banachscher Algebren

1. *Der Körper* $\boldsymbol{C}$. Die komplexen Zahlen $\{z\}$ mit der Norm

$$\|z\| = |z| = \sqrt{x^2 + y^2} \qquad (z = x + iy)$$

sind das einfachste Beispiel einer Banachschen Algebra. Die komplexen Zahlen bezeichnet man als den Körper $\boldsymbol{C}$. Für alle Elemente aus $\boldsymbol{C}$ mit Ausnahme des Nullelements ist eine Division als Umkehrung der Multiplikation erklärt. Wir werden später zeigen, daß $\boldsymbol{C}$ die einzige normierte Algebra ist, die zugleich auch ein Körper ist.

2. *Die Algebra* C_T. Es sei T ein kompakter Hausdorffraum. Mit C_T wird der lineare Raum aller auf T erklärten stetigen komplexen Funktionen bezeichnet. Addition und Multiplikation mit Zahlen sind in der für Funktionen üblichen Weise erklärt. Die Norm ist durch

$$\|x\| = \max_{t \in T} |x(t)|$$

gegeben. In den Kapiteln 2 und 3 haben wir einen Spezialfall des Raumes C_T betrachtet. $T = [a, b]$ war dort ein Intervall auf der Zahlengeraden. Ein weiteres wichtiges Beispiel eines Raumes C_T ist der Raum $C^n = \{(z_1, \ldots, z_n)\}$ der n-dimensionalen komplexen Vektoren, also der Funktionen über einem Raum aus n Punkten. Addition, Multiplikation mit Zahlen und Multiplikation der Elemente aus C^n werden koordinatenweise durchgeführt. Die Norm ist

$$\|z\| = \max_{1 \leq i \leq n} |z_i|.$$

Die Algebra C_T ist eine kommutative Banachsche Algebra. Als Einselement dient die Funktion $e(t) \equiv 1$. Das Nachprüfen der Axiome stellt keine Schwierigkeit dar.

3. *Die Algebra $\mathscr{A}$ der analytischen Funktionen im Kreis.* Mit $\mathscr{A}$ bezeichnen wir den linearen Raum aller Funktionen $x(z)$ der komplexen Veränderlichen z, die im Kreis $K = \{z: |z| \leq 1\}$ stetig und im Innern des Kreises analytisch sind. Die Multiplikation in $\mathscr{A}$ wird als die übliche Multiplikation von Funktionen erlärt. Als Norm wird

$$\|x\| = \max_{|z| \leq 1} |x(z)|$$

verwendet. Auf diesem Wege verwandeln wir $\mathscr{A}$ in eine kommutative Banachsche Algebra mit Einselement. Die Gültigkeit aller Axiome ist klar.

4. *Die Algebra l_1.* Mit l_1 wird die Gesamtheit der absolut konvergenten komplexen Zahlenfolgen

$$x = (\ldots, x_{-n}, \ldots, x_{-1}, x_0, x_1, \ldots, x_n, \ldots)$$

mit der Norm

$$\|x\| = \sum_{k=-\infty}^{\infty} |x_k| \tag{4}$$

bezeichnet. Das Produkt $x \cdot y$ der Folgen

$$x = (\ldots, x_{-n}, \ldots, x_0, \ldots, x_n, \ldots),$$
$$y = (\ldots, y_{-n}, \ldots, y_0, \ldots, y_n, \ldots)$$

wird Faltung $z = x * y$ genannt: Die Komponenten von z lassen sich aus

$$z_n = (x * y)_n = \sum_{k=-\infty}^{\infty} x_{n-k} y_k \tag{5}$$

bestimmen. Ordnet man jeder Folge x aus l_1 die trigonometrische Reihe

$$x(t) = \sum_{k=-\infty}^{\infty} x_k e^{ikt} \qquad (0 \leq t \leq 2\pi)$$

und analog y die Reihe $y(t)$ zu, so entsprechen die Elemente z_n aus (5) der Produktfunktion $x(t)\,y(t)$. Somit sind die Algebra l_1 und die Algebra W der absolut konvergenten Fourierreihen mit der Norm (4) isometrisch-isomorph. Unter Berücksichtigung der letzten Feststellungen prüft man leicht für l_1 die Axiome für eine Algebra und die Axiome für einen normierten Raum nach. Prüfen wir z. B. das Axiom 7. Für $z = x * y$ ergibt sich

$$\|z\| = \sum_n |z_n| = \sum_n \left| \sum_k x_{n-k} y_k \right| \leqq \sum_n \sum_k |x_{n-k}| \cdot |y_k|$$
$$\leqq \sum_k \left(\sum_n |x_{n-k}| \right) |y_k| = \|x\| \cdot \|y\|.$$

Da die Algebra W trivialerweise kommutativ ist, ist auch die Algebra l_1 kommutativ. Als Einselement in l_1 dient die Folge, die der Funktion $e(t) \equiv 1$ entspricht. Die Komponenten dieser Folge sind Null mit Ausnahme des Elements mit der Nummer Null, das den Wert 1 hat.

5. *Die Banachsche Algebra der beschränkten Operatoren.* X sei ein Banachraum. Wir betrachten den Raum $\mathscr{L}(X, X)$ aller linearen stetigen Operatoren, die X in sich abbilden. Hierbei gelten die üblichen Rechenregeln für die Addition, die Multiplikation mit Zahlen und die Multiplikation von Operatoren (vgl. 4.5.1. bis 4.5.3.). Der identische Operator dient als Einselement in $\mathscr{L}(X, X)$. Verwendet man wie üblich

$$\|A\| = \sup_{\|x\| \leqq 1} \|Ax\|$$

als Norm, so wird $\mathscr{L}(X, X)$ eine Banachsche Algebra. Axiom 7 wurde bereits früher nachgeprüft (vgl. Formel (4) aus 4.5.3.). Die Vollständigkeit von $\mathscr{L}(X, X)$ wurde als Aufgabe in 4.5.3. gestellt. $\mathscr{L}(X, X)$ ist eines der wichtigsten Beispiele für eine nichtkommutative Banachsche Algebra mit Einselement.

A.1.3. Maximale Ideale

Definition 4. Ein *Ideal* I einer kommutativen Algebra X ist ein solcher Unterraum von X, daß für alle $y \in I$ und alle $x \in X$ das Produkt yx zu I gehört. Das Ideal, das nur aus dem Nullelement besteht, und das Ideal, das mit X übereinstimmt, werden als triviale Ideale bezeichnet und aus den weiteren Betrachtungen ausgeschlossen. Ein Ideal heißt *maximal*, falls es in keinem anderen nichttrivialen Ideal enthalten ist.

Die eingeführten Begriffe erläutern wir am Beispiel der Algebra C_T. Es sei $\mathscr{F}$ eine nichtleere Untermenge der kompakten Menge T. Wie man leicht sieht, ist die Menge

$$M_{\mathscr{F}} = \{x(t) \in C_T : x(t) = 0,\ t \in \mathscr{F}\}$$

der Funktionen, die auf $\mathscr{F}$ verschwinden, ein Ideal in C_T. Die Menge der maximalen Ideale erlaubt eine einfache Beschreibung, die zugleich auch ein Hinweis auf das Wesen der Theorie der kommutativen Banachschen Algebren ist.

Lemma 1. *Ein maximales Ideal der Algebra C_T stimmt mit der Gesamtheit der Funktionen aus C_T überein, die an einer fixierten (aber sonst beliebigen) Stelle τ_0 der Menge T verschwinden.*

Beweis. (a) Es sei $M_{\tau_0} = \{x(t) \in C_T : x(\tau_0) = 0\}$. Dann ist M_{τ_0} ein Ideal. Wir zeigen, daß es maximal ist. Es sei $x_0(t) \notin M_{\tau_0}$, also $x_0(\tau_0) \neq 0$. Für ein beliebiges $y(t) \in C_T$ setzen wir

$$z(t) = y(t) - \frac{y(\tau_0)\, x_0(t)}{x_0(\tau_0)},$$

Dann ist $z(\tau_0) = 0$ und somit $z(t) \in M_{\tau_0}$. Also ist jedes Ideal, das M_{τ_0} und ein nicht in M_{τ_0} liegendes Element (etwa $x_0(t)$) umfaßt, trivial, da es mit C_T übereinstimmt. Also ist M_{τ_0} maximal.

(b) Wir nehmen jetzt an, daß M ein beliebiges maximales Ideal ist, und zeigen, daß die zu ihm gehörenden Funktionen in einem Punkt verschwinden. Falls das nicht so ist, findet man für jeden Punkt $\tau \in T$ eine Funktion $x_\tau(t) \in M$ mit $x_\tau(\tau) \neq 0$. Da $x_\tau(t)$ stetig ist, gibt es dann eine Umgebung U_τ von τ mit $x_\tau(t) \neq 0$ für $t \in U_\tau$. Aus $T = \bigcup_\tau U_\tau$ kann man eine endliche Überdeckung $T = U_{\tau_1} \cup \cdots \cup U_{\tau_n}$ auswählen. Aus der Idealeigenschaft folgt dann, daß

$$x_0(t) = x_{\tau_1}(t)\, \overline{x_{\tau_1}(t)} + \cdots + x_{\tau_n}(t)\, \overline{x_{\tau_n}(t)} = \sum_{k=1}^{n} |x_{\tau_k}(t)|^2$$

zu M gehört. Da $x_0(t) > 0$ auf M gilt, ist $1/x_0(t)$ stetig. Folglich gehört das Element $1 = \big(1/x_0(t)\big) \cdot x_0(t)$ zu M. Dann liegt aber jedes Element $y(t) = y(t) \cdot 1$ in M. Also ist M trivial, entgegen der Annahme, daß M maximal und somit nichttrivial ist.

Wir haben somit erhalten, daß zwischen den maximalen Idealen und den Punkten der Trägermenge T eine eineindeutige Zuordnung besteht. Man kann also die Funktionen über T als „Funktionen über der Menge der maximalen Ideale" ansehen.

Es ist eines der Ziele der hier beschriebenen Theorie der kommutativen Banachschen Algebren, zu zeigen, daß man jede derartige Algebra als Unteralgebra der Algebra der stetigen Funktionen über einem kompakten Hausdorffraum (dem Raum der maximalen Ideale) darstellen kann.

A.2. Spektrum und Resolvente

In diesem Abschnitt ist die Algebra X nicht notwendig kommutativ. Sie besitzt aber ein Einselement. Viele Betrachtungen sind analog zu den Untersuchungen aus 4.5.

A.2.1. Definitionen und Beispiele

Definition. Ein Element $x \in X$ heißt *umkehrbar* oder *invertierbar*, falls es ein Element x^{-1}, das *inverse Element* zu x, mit

$$x \cdot x^{-1} = x^{-1} \cdot x = e$$

gibt. Existiert kein derartiges Element, so heißt x *nicht umkehrbar* oder *nicht invertierbar*. Das *Spektrum* $\sigma(x)$ des Elements $x \in X$ ist die Gesamtheit der komplexen Zahlen λ, für die $(\lambda e - x)$ nicht invertierbar ist. Jede Zahl λ mit $\lambda \notin \sigma(x)$ heißt *regulär*. Die Funktion

$$R_\lambda \colon \boldsymbol{C} \setminus \sigma(x) \to X, \qquad R_\lambda x = x(\lambda) = (\lambda e - x)^{-1},$$

die auf der Menge der regulären Zahlen bezüglich des Elements x bestimmt ist, wird *Resolvente* des Elements x genannt. Als *Spektralradius* $r(x)$ des Elements $x \in X$ wird

$$r(x) = \sup_{\lambda \in \sigma(x)} |\lambda| \tag{1}$$

bezeichnet.

Die eingeführten wichtigen Begriffe erläutern wir an Beispielen.

(a) Ist $X = \boldsymbol{C}$, so sind alle von Null verschiedenen Elemente invertierbar.

(b) Ist $X = C_T$, so ist $x(t)$ genau dann invertierbar, wenn $x(t)$ überall von Null verschieden ist. Das Spektrum $\sigma(x)$ besteht aus dem Wertevorrat von $x(t)$. Die Resolvente hat die Form

$$R_\lambda = \frac{1}{\lambda - x(t)}.$$

Ferner ist

$$r(x) = \|x\| = \max_t |x(t)|$$

(c) Ist $X = \mathscr{L}(Y, Y)$ die Algebra der beschränkten Operatoren, so sind die inversen Elemente die inversen Operatoren. Spektrum und Resolvente stimmen in diesem Fall mit dem Spektrum und der Resolvente des Operators überein (vgl. 4.5.7.). Wir werden in diesem Abschnitt jene Begriffe allgemein untersuchen, die wir früher für die Banachsche Algebra der beschränkten Operatoren eingeführt haben.

A.2.2. Eigenschaften des Spektrums

Satz 1. 1. *Es sei f ein lineares stetiges Funktional aus dem adjungierten Raum X^*. Dann ist $F(\lambda) = f\big(x(\lambda)\big)$ in $\boldsymbol{C} \setminus \sigma(x)$ analytisch, und es gilt $F(\lambda) \to 0$ für $|\lambda| \to \infty$.*

2. *Das Spektrum* $\sigma(x)$ *eines Elements* x *der Banachschen Algebra* X *ist eine nichtleere kompakte Menge in* $\mathbf{C}$. *Es gilt*

$$r(x) \leqq \|x\|. \tag{2}$$

Zum Beweis des Satzes benötigen wir einige Lemmata.

Lemma 1 (vgl. Satz 5 aus 4.5.). *Es sei* x *ein Element aus der Banachschen Algebra* X *mit* $\|x\| < 1$. *Dann ist* $e - x$ *invertierbar, und es gilt*

$$(e - x)^{-1} = e + x + \cdots + x^n + \cdots.$$

Beweis. Mit $s_n = e + x + \cdots + x^n$ erhält man

$$\|s_n - s_{n+k}\| = \|x^{n+1} + \cdots + x^{n+k}\| \leqq \|x^{n+1}\| \sum_{i=0}^{k-1} \|x\|^i \leqq \|x\|^{n+1} \frac{1}{1 - \|x\|} \to 0$$

für $n \to \infty$. Somit ist $\{s_n\}$ eine Fundamentalfolge. Da X vollständig ist, konvergiert s_n gegen ein Element $s \in X$. Es gilt

$$s(e - x) = \lim_{n\to\infty} s_n(e - x) = \lim_{n\to\infty} (e - x^{n+1}) = e.$$

Entsprechend beweist man $(e - x)\, s = e$.

Folgerung. *Für jedes* $x \in X$ *gilt*

$$(e - tx)^{-1} \to e \quad \textit{für } t \to 0.$$

Diese Aussage ergibt sich aus

$$(e - tx)^{-1} = \lim_{n\to\infty} (e + tx + \cdots + (tx)^n) = e + O(t).$$

Lemma 2 (vgl. Satz 4 aus 4.5.). *Es sei* x_0 *ein invertierbares Element, und ferner sei* $\|\Delta x\| < \|x_0^{-1}\|^{-1}$. *Dann ist* $x_1 = x_0 + \Delta x$ *invertierbar, und es gilt*

$$x_1^{-1} = (e - x_0^{-1}\,\Delta x)^{-1}\, x_0^{-1}.$$

Beweis. Es ist

$$x_1 = x_0 + \Delta x = x_0(e + x_0^{-1}\,\Delta x) = x_0(e - x)$$

und $\|x\| = \|-x_0^{-1}\,\Delta x\| < 1$. Nach Lemma 1 ist $x_1^{-1} = (e - x)^{-1}\, x_0^{-1}$, was zu zeigen war.

Folgerung 1. *Die Menge der invertierbaren Elemente einer Banachschen Algebra ist offen (in der Normtopologie der Banachschen Algebra). Die Menge der nichtinvertierbaren Elemente ist abgeschlossen.*

Folgerung 2. *Die Resolvente* $x(\lambda)$ *ist auf* $\mathbf{C} \setminus \sigma(x)$ *eine stetige Funktion von* λ.

Nach Lemma 2 und der Folgerung aus Lemma 1 ist nämlich

$$x(\lambda_0 + \Delta\lambda) = (\lambda_0 e - x + \Delta\lambda e)^{-1} = \big(e + \Delta\lambda x\,(\lambda_0)\big)^{-1}\, x(\lambda_0) \to x(\lambda_0)$$

für $\Delta\lambda \to 0$.

Lemma 3 (vgl. 4.5.7.). *Für* $\lambda, \mu \in \boldsymbol{C} \setminus \sigma(x)$ *ist*

(a) $R_\lambda x \cdot R_\mu x = R_\mu x \cdot R_\lambda x$,

(b) $R_\lambda x - R_\mu x = (\mu - \lambda)\, R_\lambda x \cdot R_\mu x$ (*Hilbertsche Relation*).

Beweis. (a) Es ist

$$\begin{aligned} R_\lambda x \cdot R_\mu x &= (\lambda e - x)^{-1} (\mu e - x)^{-1} = [(\mu e - x)(\lambda e - x)]^{-1} \\ &= [(\lambda e - x)(\mu e - x)]^{-1} = (\mu e - x)^{-1} (\lambda e - x)^{-1} = R_\mu x \cdot R_\lambda x. \end{aligned}$$

(b) Unter Verwendung von (a) erhält man aus der Definition von R_λ und R_μ

$$R_\lambda x = (\mu e - x)\, R_\lambda x \cdot R_\mu x,$$

$$R_\mu x = (\lambda e - x)\, R_\lambda x \cdot R_\mu x.$$

Hieraus ergibt sich

$$R_\lambda x - R_\mu x = (\mu e - \lambda e)\, R_\lambda x \cdot R_\mu x = (\mu - \lambda)\, R_\lambda x \cdot R_\mu x,$$

was zu zeigen war.

Folgerung. *Ist* $\lambda_0 \in \boldsymbol{C} \setminus \sigma(x)$, *so gilt* $x'(\lambda_0) = -x^2(\lambda_0)$.

Aus (b) und aus der Folgerung 2 zu Lemma 2 erhält man nämlich

$$x'(\lambda_0) = \lim_{\lambda \to \lambda_0} \frac{x(\lambda) - x(\lambda_0)}{\lambda - \lambda_0} = -\lim_{\lambda \to \lambda_0} x(\lambda)\, x(\lambda_0) = -x^2(\lambda_0).$$

Beweis von Satz 1.

1. Es sei $f \in X^*$ ein lineares stetiges Funktional über X. Wir setzen

$$F(\lambda) = f\big(x(\lambda)\big) = f(R_\lambda x).$$

Für $\lambda_0 \notin \sigma(x)$ ergibt sich aus der Folgerung zu Lemma 3

$$\begin{aligned} F'(\lambda_0) &= \lim_{\lambda \to \lambda_0} \frac{F(\lambda) - F(\lambda_0)}{\lambda - \lambda_0} = \lim_{\lambda \to \lambda_0} f\left(\frac{x(\lambda) - x(\lambda_0)}{\lambda - \lambda_0}\right) \\ &= f\left(\lim_{\lambda \to \lambda_0} \frac{x(\lambda) - x(\lambda_0)}{\lambda - \lambda_0}\right) = -f\big(x^2(\lambda_0)\big). \end{aligned}$$

Hieraus folgt, daß $F(\lambda)$ analytisch ist. Für $|\lambda| > \|x\|$ erhält man ferner aus Lemma 1

$$\begin{aligned} |F(\lambda)| &\leqq \|f\|_{X^*} \|x(\lambda)\|_X = \|f\|_{X^*} \|(\lambda e - x)^{-1}\| \\ &= \frac{\|f\|_{X^*}}{|\lambda|} \left\| \left(e - \frac{x}{\lambda}\right)^{-1} \right\| \leqq \frac{C}{|\lambda|} \to 0 \quad \text{für } \lambda \to \infty. \end{aligned}$$

2. (a) Wir beweisen, daß das Spektrum $\sigma(x)$ nicht leer ist. Es sei $\sigma(x) = \emptyset$. Aus Teil 1 des Satzes folgt, daß $F(\lambda)$ für $f \in X^*$ eine ganze analytische Funktion ist, die für $|\lambda| \to \infty$ gegen 0 strebt. Nach dem Satz von Liouville aus der komplexen Funk-

tionentheorie ist somit $F(\lambda) \equiv 0$. Also ist $f(x^{-1}) = -f\big((0e - x)^{-1}\big) = 0$ für jedes $f \in X^*$. Aus dem Satz von HAHN-BANACH (vgl. 4.1.3.) ergibt sich dann $x^{-1} = 0$. Das ist ein Widerspruch.

(b) Wir beweisen, daß $\sigma(x)$ kompakt ist. Nach Lemma 1 ist

$$\lambda e - x = \lambda \left(e - \frac{x}{\lambda}\right) \quad \text{für} \quad |\lambda| > \|x\|$$

invertierbar. Somit ist $\sigma(x)$ beschränkt, und es gilt Formel (2). Die Abgeschlossenheit von $\sigma(x)$ folgt leicht aus Lemma 2: Ist λ_0 regulär, so sind auch die Punkte $\lambda_0 + \Delta\lambda$ mit $|\Delta\lambda| < \|x(\lambda_0)\|$ regulär, da

$$(\lambda_0 + \Delta\lambda)\, e - x = \lambda_0 e - x + \Delta\lambda e$$

gilt.

Wir notieren zwei Folgerungen aus Satz 1.

Folgerung 1. *Jede Banachsche Algebra (über $\boldsymbol{C}$), die zugleich ein Körper ist, ist isometrisch-isomorph zu $\boldsymbol{C}$.*

Beweis. Es sei X ein „Banachscher Körper" und x ein beliebiges Element aus X. Ist λ eine Zahl, so daß $\lambda e - x$ nicht invertierbar ist, so muß $\lambda e - x = 0$ sein. Also ist $x = \lambda e$. Man sieht leicht, daß die Zuordnung $\lambda \leftrightarrow x$ ein Isomorphismus von X auf $\boldsymbol{C}$ ist. Da $\|e\| = 1$ gilt, muß $\|x\| = |\lambda|$ sein. Somit sind X und $\boldsymbol{C}$ isometrisch.

Folgerung 2. *Das Spektrum eines beliebigen Operators A aus $\mathscr{L}(X, X)$ ist nicht leer.*

Diese Aussage wurde bereits füher ohne Beweis formuliert. (Man vergleiche die Bemerkungen am Ende von 4.5.7.)

A.2.3. Der Satz über den Spektralradius

Satz 2. *Den Spektralradius kann man mit Hilfe der Formel*

$$r(x) = \lim_{n\to\infty} \sqrt[n]{\|x^n\|} \tag{3}$$

berechnen.

Beweis. Ist f ein Funktional aus X^*, so folgt aus Satz 1, daß $F(\lambda) = f\big(x(\lambda)\big)$ in $\boldsymbol{C} \setminus \sigma(x)$ analytisch ist. Insbesondere ist $F(\lambda)$ im Gebiet $|\lambda| > \|x\|$ analytisch. Für derartige Werte von λ ergibt sich aber aus Lemma 1

$$x(\lambda) = (\lambda e - x)^{-1} = \frac{1}{\lambda}\left(e - \frac{x}{\lambda}\right)^{-1} = \sum_{n=0}^{\infty} \frac{x^n}{\lambda^{n+1}}.$$

Somit ist

$$F(\lambda) = f\big(x(\lambda)\big) = \sum_{n=0}^{\infty} \frac{f(x^n)}{\lambda^{n+1}}.$$

Aus dem Eindeutigkeitssatz für analytische Funktionen folgt aber, daß diese Darstellung nicht nur für $|\lambda| > \|x\|$, sondern für alle λ mit $|\lambda| > r(x)$ gilt. Insbesondere ist dann

$$\sup_n \left|\frac{f(x^n)}{\lambda^{n+1}}\right| < \infty.$$

Also sind die Elemente $\dfrac{x^n}{\lambda^{n+1}}$ schwach beschränkt. Somit sind sie auch stark beschränkt. (Dieses Resultat, das Prinzip der gleichmäßigen Beschränktheit oder der Satz von BANACH-STEINHAUS, wurde in der Bemerkung in 4.3.2. bewiesen. Man vergleiche hierzu auch [12], Kap. II.) Folglich existiert eine von λ abhängige Zahl $c(\lambda)$ mit

$$\left\|\frac{x^n}{\lambda^{n+1}}\right\| \leqq c(\lambda).$$

Hieraus ergibt sich

$$\overline{\lim_{n\to\infty}} \|x^n\|^{1/n} \leqq |\lambda| \qquad \text{für alle } \{\lambda\colon |\lambda| > r(x)\}.$$

Somit ist

$$\overline{\lim_{n\to\infty}} \|x^n\|^{1/n} \leqq r(x).$$

Es sei jetzt $\lambda \in \sigma(x)$. Dann ist $\lambda^n \in \sigma(x^n)$, da $\lambda^n e - x^n$ das Element $\lambda e - x$ als Faktor enthält. Für $\mu \in \sigma(x)$ ergibt sich $|\mu| \leqq \|x\|$ (Satz 1). Wählt man $\mu = \lambda^n$ mit $\lambda \in \sigma(x)$, so erhält man $|\lambda| \leqq \sqrt[n]{\|x^n\|}$. Somit ist

$$r(x) \leqq \underline{\lim_{n\to\infty}} \sqrt[n]{\|x^n\|}.$$

Hieraus folgt der Satz.

A.3. Vorbereitende Betrachtungen

In diesem kurzen Abschnitt leiten wir einige vorbereitende Aussagen mit Hilfe der üblichen Beweistechnik her.

A.3.1. Der Satz über die Faktoralgebra. Es sei X eine kommutative Banachsche Algebra mit Einselement. Ferner sei I ein Ideal in X. Wir stellen zuerst fest, daß I nur aus nicht invertierbaren Elementen besteht. Wäre nämlich $z \in I$ invertierbar, so würde jedes Element $x \in X$ ebenfals zu I gehören, da $x = (xz^{-1})\, z \in I$ gilt. Somit

wäre I trivial, was aber ausgeschlossen wurde. Ferner bemerken wir, daß nach Lemma 1 aus A.2.2. der Abstand des Einselements e zu einem beliebigen nicht invertierbaren Element (und somit auch zu einem beliebigen Ideal) mindestens gleich 1 ist.

Wir betrachten jetzt den Faktorraum X/I (vgl. 3.1.4.) und definieren in ihm eine Multiplikation. Ist x aus der Klasse $\xi \in X/I$ und y aus der Klasse $\eta \in X/I$, so wird das Produkt als die Klasse ζ erklärt, in der $x \cdot y$ liegt. (Man prüfe nach, daß das Resultat unabhängig von der Wahl der Repräsentanten x und y aus den Klassen ξ und η ist. Ferner zeige man, daß die so erklärte Multiplikation die Axiome 1 bis 5 aus A.1.1. erfüllt). Damit ist X/I eine kommutative Algebra. Sie wird als *Faktoralgebra* von X nach dem Ideal I bezeichnet.

Durch

$$\|\xi\| = \inf_{y \in I} \|x + y\|$$

wird in X/I eine Norm eingeführt. Hierbei ist x ein Repräsentant der Klasse ξ.

Satz 1. *Ist X eine Banachsche Algebra und I ein abgeschlossenes Ideal in X, so ist die Faktoralgebra X/I ebenfalls eine Banachsche Algebra.*

Beweis. Wir müssen zeigen, daß erstens $\|\xi\|$ eine Norm ist und daß zweitens X/I bezüglich dieser Norm vollständig ist.

1. (a) Es gilt $\|\xi\| = \inf_{y \in I} \|x + y\| \geqq 0$. Ferner ist $\|0\| = \inf_{y \in I} \|0 + y\| = 0$. Es sei $\|\xi_0\| = 0$. Dann folgt aus $\inf_{y \in I} \|x_0 + y\| = 0$, $x_0 \in \xi_0$, die Existenz einer Folge $y_n \in I$ mit $\|x_0 + y_n\| \to 0$, also $(-y_n) \to x_0$. Da I abgeschlossen ist, gehört x_0 zu I. Somit ist $\xi_0 \sim 0$. Damit ist $\|\xi\| \geqq 0$ gezeigt, wobei $\|\xi\| = 0$ genau dann gilt, wenn $\xi = 0$ ist.

(b) Für $\lambda \neq 0$ ist

$$\|\lambda\xi\| = \inf_{y \in I} \|\lambda x + y\| = |\lambda| \inf_{y \in I} \left\| x + \frac{y}{\lambda} \right\| = |\lambda| \cdot \|\xi\|.$$

Für $\lambda = 0$ ist $\|\lambda\xi\| = 0 = |\lambda| \cdot \|\xi\|$.

(c)
$$\begin{aligned} \|\xi + \eta\| &= \inf_{z \in I} \|x + y + z\| = \inf_{u,v \in I} \|x + y + u + v\| \\ &\leqq \inf_{u \in I} \|x + u\| + \inf_{v \in I} \|y + v\| = \|\xi\| + \|\eta\|. \end{aligned}$$

(d)
$$\begin{aligned} \|\xi\eta\| &= \inf_{z \in I} \|xy + z\| \leqq \inf_{u,v \in I} \|(x + u)(y + v)\| \\ &\leqq \inf_{u \in I} \|x + u\| \cdot \inf_{v \in I} \|y + v\| = \|\xi\| \cdot \|\eta\|. \end{aligned}$$

(e) Aus $E = e + I$ ergibt sich $E^2 = e^2 + I = e + I = E$. Somit ist $\|E\| = \|E^2\| \leqq \|E\|^2$. Wie wir oben bemerkt haben, enthält eine Umgebung von e keine nicht invertierbaren Elemente. Somit ist E nicht das Nullelement. Dann gilt $1 \leqq \|E\|$. Da

andererseits

$$\|E\| = \inf_{y \in I} \|e + y\| \leqq 1$$

ist, finden wir $\|E\| = 1$.

2. Wir zeigen nun, daß X/I vollständig ist. $\xi_1, \ldots, \xi_n, \ldots$ sei eine Fundamentalfolge, also $\|\xi_{n+m} - \xi_n\| < \varepsilon$ für $n > N(\varepsilon)$ und $m \geqq 1$. Wir wählen n_k so, daß $\|\xi_{n_k+m} - \xi_{n_k}\| \leqq 1/2^k$ gilt. Insbesondere ist $\|\xi_{n_2} - \xi_{n_1}\| \leqq 1/2$. Für zwei geeignete Repräsentanten $x_1 \in \xi_{n_1}$ und $x_2 \in \xi_{n_2}$ erhalten wir dann $\|x_2 - x_1\| \leqq 1$. Analog konstruieren wir $x_k \in \xi_{n_k}$ $(k = 2, 3, 4, \ldots)$ mit $\|x_{k+1} - x_k\| \leqq 1/2^{k-1}$. Dann ist $\{x_k\}$ eine Fundamentalfolge in X. Da X vollständig ist, existiert ein Grenzelement $x_0 = \lim\limits_{k \to \infty} x_k$. Ist $\xi_0 = x_0 + I$, so folgt

$$\|\xi_{n_k} - \xi_0\| = \inf_{y \in I} \|x_k - x_0 + y\| \leqq \|x_k - x_0\| \to 0 \quad \text{für } k \to \infty.$$

Somit ist $\xi_{n_k} \to \xi_0$. Also ist X/I vollständig. Damit ist der Satz bewiesen.

A.3.2. Drei Lemmata. Wir benötigen später drei Lemmata, ein mengentheoretisches, ein algebraisches und ein topologisches.

Lemma 1. *Jedes nichttriviale Ideal ist in einem maximalen Ideal enthalten.*

Beweis. Der Beweis stützt sich auf das Lemma von Zorn, das in 1.4.7. formuliert wurde. $\mathscr{I}$ sei die Menge aller Ideale, die ein vorgegebenes Ideal I umfassen. Bezüglich der mengentheoretischen Inklusion ist $\mathscr{I}$ halbgeordnet: $I_1 \leqq I_2$, falls $I_1 \subset I_2$. Für jede linear geordnete Menge $\{I_\alpha\}$ aus $\mathscr{I}$ ist $\bigcup\limits_\alpha I_\alpha$ ein nichttriviales Ideal, das als obere Schranke von $\{I_\alpha\}$ dient. Aus dem Lemma von Zorn folgt dann, daß I in einem maximalen Ideal enthalten ist.

Folgerung. *Ist X kein Körper, so gibt es in ihm maximale Ideale. Jedes von 0 verschiedene nicht invertierbare Element ist in einem maximalen Ideal enthalten.*

Ist $x_0 \neq 0$ ein nicht invertierbares Element, so ist x_0X ein Ideal. Es enthält x_0, aber nicht e und ist somit nicht trivial. Nach Lemma 1 ist es in einem passenden maximalen Ideal enthalten.

Lemma 2. *Ein Ideal I liegt genau dann echt in einem nichttrivialen Ideal $I' \subset X$, wenn die Faktoralgebra X/I nichttriviale Ideale besitzt.*

Beweis. Wir beweisen die Notwendigkeit. Es sei $I \subset I' \subset X$, $I \neq I'$ und $X \neq I'$. In der Klasse $\xi = x + I$ betrachten wir die Unterklasse derjenigen ξ', für die x' zu I' gehört. Man prüft leicht nach, daß man ein nichttriviales Ideal in X/I enthält. Analog beweist man die Hinlänglichkeit.

Lemma 3. *Der Abschluß eines Ideals I ist ein (nichttriviales) Ideal.*

Beweis. Der Abschluß ist nichttrivial, da I nur aus nicht invertierbaren Elementen besteht und die nicht invertierbaren Elemente eine abgeschlossene Menge sind. Alles andere folgt aus der Stetigkeit der algebraischen Operationen.

Folgerung. *Maximale Ideale sind abgeschlossen.*

A.4. Hauptsätze

In diesem Abschnitt ist X eine kommutative Banachsche Algebra mit Einselement.

A.4.1. Lineare stetige multiplikative Funktionale und maximale Ideale

Definition 1. Ein lineares stetiges Funktional f auf einer Banachschen Algebra X nennt man *multiplikativ*, falls für beliebige x und y

$$f(x \cdot y) = f(x) \cdot f(y) \tag{1}$$

gilt. Die Gesamtheit der nichttrivialen[1]) linearen multiplikativen Funktionale bezeichnen wir mit $\mathscr{M}$.

Wir bemerken, daß wir ein lineares stetiges multiplikatives Funktional auch als stetigen Homomorphismus von X in $\boldsymbol{C}$ definieren können. Ist $f \in \mathscr{M}$, so gilt

$$|f(x)| \leqq \|x\|. \tag{2}$$

Um das zu zeigen, nehmen wir an, daß $|f(x_0)| = \lambda > 1$ für ein Element x_0 mit $\|x_0\| = 1$ gilt. Dann erhält man aus

$$|f(x_0{}^n)| = \lambda^n \to \infty$$

einen Widerspruch zur Beschränktheit (Stetigkeit) von f. Ferner ist

$$f(e) = f(e^2) = [f(e)]^2.$$

Also ist entweder $f(e) = 0$ oder

$$f(e) = 1. \tag{3}$$

Im ersten Fall ist f trivial. Aus (2) und (3) folgt, daß nichttriviale lineare stetige multiplikative Funktionale die Norm 1 haben. Somit ist $\mathscr{M}$ eine Untermenge der Oberfläche der Einheitskugel im adjungierten Raum X^*.

Mit $\operatorname{Ker} f = \{x : f(x) = 0\}$ wird der Nullraum des Funktionals f bezeichnet. $\operatorname{Ker} f$ heißt *Kern* von f.

Lemma 1. *Für $f \in \mathscr{M}$ ist der Kern* Ker f *ein maximales Ideal.*

---

[1]) Anm. d. Übers.: Hierdurch wird das Funktional $f(x) = 0$ für alle $x \in X$ ausgesondert.

Beweis. Aus $y \in I = \operatorname{Ker} f$ und $x \in X$ folgt

$$f(y \cdot x) = f(y) \cdot f(x) = 0, \quad \text{d. h. } y \cdot x \in \operatorname{Ker} f.$$

Also ist $\operatorname{Ker} f$ ein Ideal. Um zu zeigen, daß $\operatorname{Ker} f$ maximal ist, betrachten wir ein Ideal $I \neq X$ mit $\operatorname{Ker} f \subset I$. Es sei $x_0 \in I$ und $x_0 \notin \operatorname{Ker} f$. Da $\operatorname{Ker} f$ die Kodimension 1 hat (vgl. 3.1.6.) und abgeschlossen ist, kann man e in der Form

$$e = \lambda x_0 + y, \quad y \in \operatorname{Ker} f,$$

darstellen. Hieraus folgt $e \in I$. Somit ist $I = X$. Das ist ein Widerspruch.

Lemma 2. *Zu jedem maximalen Ideal M gehört ein eindeutig bestimmtes lineares stetiges multiplikatives Funktional $f \in \mathcal{M}$ mit $M = \operatorname{Ker} f$.*

Beweis. Die Folgerung zu Lemma 3 aus A.3.2. zeigt, daß M ein abgeschlossenes Ideal ist. Aus Satz 1 aus A.3.1. folgt, daß X/M eine Banachsche Algebra ist. Aus Lemma 2 aus A.3.2. erhält man, daß X/M keine nichttrivialen Ideale und somit auch keine von 0 verschiedenen nicht invertierbaren Elemente enthält (vgl. auch die Folgerung zu Lemma 1 aus A.3.2.). Somit ist die Banachsche Algebra X/M zugleich ein Körper. Folgerung 1 zu Satz 1 aus A.2.2. zeigt dann, daß X/M isomorph zu $\boldsymbol{C}$ ist. Das bedeutet, daß man jedes $x \in X$ mit Hilfe einer eindeutig bestimmten Zahl $f(x) \in \boldsymbol{C}$ in der Form

$$x = f(x) \cdot e + u, \quad u \in M, \tag{4}$$

darstellen kann. Es zeigt sich, daß f ein Homomorphismus ist. Wir beweisen z. B. $f(x \cdot y) = f(x) \cdot f(y)$. Aus

$$x = f(x) \cdot e + u, \quad u \in M,$$

$$y = f(y) \cdot e + v, \quad v \in M,$$

folgt $x \cdot y = f(x) \cdot f(y) \cdot e + w$ mit $w \in M$. Somit ist $f(x \cdot y) = f(x) \cdot f(y)$. Analog läßt sich $f(x + y) = f(x) + f(y)$ und $f(\lambda x) = \lambda f(x)$ beweisen. Aus (4) folgt ferner $f(x) = 0$ für $x \in M$ und $f(e) = 1$. Damit ist das Lemma bewiesen.

Wir haben somit eine eineindeutige Zuordnung zwischen den maximalen Idealen $\{M\}$ und den Funktionalen f aus $\mathcal{M}$ erhalten. Wir bezeichnen deshalb die Funktionale aus $\mathcal{M}$ mit f_M, wobei M das zugehörige maximale Ideal ist. Für die Menge der maximalen Ideale $\{M\}$ sowie für die zugehörige Menge $\{f_M\}$ benutzen wir den Buchstaben $\mathcal{M}$.

Für $x \in X$ betrachten wir auf $\mathcal{M}$ die Funktion $x(M)$, die durch

$$x(M) = f_M(x) \tag{5}$$

definiert ist. (Der Wert der Funktion $x(M) = f_M(x)$ wird durch den Homomorphismus ermittelt, der durch (4) beschrieben wird.) Wir haben somit eine Darstellung der Elemente der Algebra X in Form von Funktionen über $\mathcal{M}$ erhalten. (Man vergleiche hierzu das am Schluß von A.1. Gesagte.)

A.4.2. Die Topologie auf der Menge $\mathscr{M}$. Hauptsätze. Wir müssen noch zeigen, daß $\mathscr{M}$ in einer gewissen Topologie kompakt ist und daß die Funktionen $x(M)$ bezüglich dieser Topologie stetig sind.

Wir hatten schon erwähnt, daß $\mathscr{M}$ eine Untermenge der Oberfläche der Einheitskugel in X^* ist. In 4.3.4. hatten wir andererseits für den separablen Fall folgende Aussage bewiesen:

Die Einheitskugel im adjungierten Raum X^* zum Banachraum X ist w^*-kompakt.

Einen Beweis dieses Satzes für den allgemeinen Fall findet man z. B. in [12], V.4. Wir erinnern daran, daß die w^*-Topologie durch das System der Umgebungen

$$U_{x_1,\ldots,x_m,\delta}(f_0) = \{f \in X^*\colon |f(x_k) - f_0(x_k)| < \delta,\ k = 1, \ldots, m\} \tag{6}$$

bestimmt wurde. Die Menge $\mathscr{M}$ versehen wir jetzt mit dieser w^*-Topologie. Die Kompaktheit von $\mathscr{M}$ folgt aus dem oben angegebenen Resultat und dem folgenden Lemma.

Lemma 3. *Die Menge $\mathscr{M}$ ist eine abgeschlossene Untermenge der Oberfläche der Einheitskugel in X^*. Die Funktionen $x(M)$ sind auf $\mathscr{M}$ stetig.*

Beweis. Es sei f_0 ein Funktional aus dem Abschluß von $\mathscr{M}$, d. h., daß man in jeder Umgebung der Abbildung f_0 einen Homomorphismus f_M findet, der durch ein maximales Ideal M erzeugt wird. Wir wählen die Umgebung $U_{x,y,x+y,\delta}(f_0)$. Aus (6) und aus der Definition von $x(M)$ folgt

$$\left.\begin{aligned} &|f_M(x) - f_0(x)| < \delta, \\ &|f_M(y) - f_0(y)| < \delta, \\ &|f_M(x+y) - f_0(x+y)| < \delta. \end{aligned}\right\} \tag{7}$$

Da f_M ein Homomorphismus ist, gilt $f_M(x + y) = f_M(x) + f_M(y)$. Dann ergibt sich aber aus (7)

$$f_0(x + y) = f_0(x) + f_0(y).$$

Analog beweist man $f_0(\alpha x) = \alpha f_0(x)$ und $f_0(xy) = f_0(x)\, f_0(y)$. (Hierzu benötigt man die Umgebungen $U_{x,\alpha x,\delta}(f_0)$ und $U_{x,y,xy,\delta}(f_0)$.) Somit ist f ein stetiges lineares multiplikatives Funktional. Gehen wir von der Umgebung $U_{e,\delta}(f_0)$ aus, so erhalten wir $f_0(e) = 1$. Somit ist f_0 nicht trivial. Also ist $f_0 \in \mathscr{M}$. Hieraus folgt, daß $\mathscr{M}$ abgeschlossen ist.

Wir zeigen jetzt, daß $x_0(M) = f(x_0)$ auf $\mathscr{M}$ stetig ist. Es sei $M_0 \in \mathscr{M}$. Wir wählen eine Umgebung $U_{x_0,\varepsilon}(M_0)$, $\varepsilon > 0$. Ist $M \in U_{x_0,\varepsilon}(M_0)$, so folgt aus (6)

$$|f_M(x_0) - f_{M_0}(x_0)| = |x_0(M) - x_0(M_0)| < \varepsilon.$$

Das bedeutet aber, daß $x_0(M)$ im Punkt M_0 stetig ist.

Satz 1. *Die Abbildung $x \to x(M)$ ist ein Homomorphismus der Algebra X in die Algebra $C_{\mathscr{M}}$ der stetigen Funktionen über dem kompakten Hausdorffraum $\mathscr{M}$ der maxi-*

malen Ideale der Algebra X. Hierbei ist

$$\|x(M)\| = \max |x(M)| \leqq \|x\|. \tag{8}$$

Beweis. Aus dem vorher Gesagten folgt, daß wir nur noch die Formel (8) beweisen müssen. Nach Definition gehört $x - f_M(x)\, e$ zum Ideal M. Somit ist dieses Element nicht invertierbar. Also ist $f_M(x) \in \sigma(x)$. Wählt man andererseits eine beliebige Zahl $\lambda_0 \in \sigma(x)$, so ist $x - \lambda_0 e$ nicht invertierbar. Dieses Element gehört somit zu einem maximalen Ideal M, und es gilt $0 = f_M(x - \lambda_0 e)$. Somit ist $\lambda_0 = f_M(x)$. Also stimmt das Bild von $\mathcal{M}$ bei der Abbildung $x(M)$ mit $\sigma(x)$ überein. Mit Hilfe von Satz 1 aus A.2.2. ergibt sich dann (8).

Wir verschärfen jetzt Satz 1 bei Annahme zusätzlicher Eigenschaften über die Algebra X. Hierzu führen wir drei Begriffe ein.

Definition 2. Der Durchschnitt $R = \bigcap\limits_{\mathcal{M}} M$ aller maximaler Ideale heißt *Radikal* von X. Eine Banachsche Algebra X heißt *regulär*, falls $\|x^2\| = \|x\|^2$ gilt. Eine Banachsche Algebra X heißt *symmetrisch*, falls man zu jeder Funktion $x(M)$ ein Element $y \in X$ mit $y(M) = \overline{x(M)}$ findet (konjugiert-komplexe Funktion).

Satz 2. (a) *Besteht das Radikal der Algebra X nur aus dem Nullelement, so ist die Abbildung $x \to x(M)$ eineindeutig.*

(b) *Ist die Algebra X regulär, so ist X isometrisch-isomorph zu seinem Bild in $C_{\mathcal{M}}$. Insbesondere besteht das Radikal von X nur aus dem Nullelement.*

(c) *Ist die Algebra X symmetrisch, so ist das Bild von X bei der Abbildung $x \to x(M)$ dicht in $C_{\mathcal{M}}$.*

(d) *Besitzt die Algebra X die Eigenschaften* (b) *und* (c), *so ist X isometrisch-isomorph zu $C_{\mathcal{M}}$.*

Beweis. Zuerst beweisen wir die letzte Aussage aus den vorhergehenden. Nach (b) ist die eineindeutige Abbildung $x \leftrightarrow \{x(M)\}$ isometrisch, $\|x\|_X = \max\limits_{M \in \mathcal{M}} |x(M)|$. Nach (c) ist $\{x(M)\}$ dicht in $C_{\mathcal{M}}$. Da X vollständig ist, ist somit auch $\{x(M)\}$ vollständig. Hieraus folgt $\{x(M)\} = C_{\mathcal{M}}$.

Wir beweisen nun (a). Es sei $x_0 \neq 0$ und $x_0(M) \equiv 0$ auf $\mathcal{M}$. Das bedeutet $f_M(x_0) = 0$ für alle M und somit $x_0 \in M$ für alle M. Also ist $x_0 \in R$. Da $R = \{0\}$ gilt, muß $x_0 = 0$ sein. Das ist ein Widerspruch und (a) somit bewiesen.

Um (b) zu beweisen, bemerken wir, daß aus $\|x^2\| = \|x\|^2$ sofort

$$\lim_{n\to\infty} \sqrt[2^n]{\|x^{2^n}\|} = \|x\|$$

folgt. Mit Hilfe von Satz 2 aus A.2.3. ergibt sich dann für den Spektralradius

$$r(x) = \|x\|. \tag{9}$$

(9) zeigt, daß das Radikal nur aus dem Nullelement besteht. Würde nämlich $f_M(x_0) = 0$ für ein Element $0 \neq x_0 \in R$ und für alle M gelten, so würde $\sigma(x)$ nur aus dem Nullpunkt bestehen. Dem widerspricht aber $r(x_0) = \|x_0\| \neq 0$. Somit ist die Abbildung $x \leftrightarrow x(M)$ von X auf die Unteralgebra $\{x(M)\}$ in $C_{\mathscr{M}}$ ein Isomorphismus. Dieser Isomorphismus ist isometrisch. Denn aus (8) und (9) folgt

$$\|x(M)\|_{C_{\mathscr{M}}} = \max_{M \in \mathscr{M}} |x(M)| = r(x) = \|x\|.$$

Zum Beweis von (c) benötigen wir einen der bemerkenswertesten Sätze der Algebra und der Analysis, den Satz von Stone-Weierstrass, der folgendermaßen lautet:

Es sei A eine Unteralgebra der Banachschen Algebra C_T der stetigen Funktionen über der kompakten Menge T mit den Eigenschaften:

1. *Das Einselement, d. h. die Funktion $e(t) \equiv 1$, gehört zu A.*

2. *Die Algebra A trennt die Punkte von T, d. h., für beliebige Punkte $t_1 \neq t_2$ existiert eine Funktion $x(t) \in A$ mit $x(t_1) \neq x(t_2)$.*

3. *Die Algebra A ist invariant bezüglich der Konjugiertenbildung, d. h., aus $x(t) \in A$ folgt $\overline{x(t)} \in A$.*

Dann ist A dicht in C_T.

Einen Beweis des Satzes von Stone-Weierstrass findet man in [17], S. 53–56, [12], IV.6, und [70], Einleitung, Abschnitt 2.

Wir beweisen jetzt (c). Es sei $A = \{x(M)\}$ das Bild von X bei der Abbildung $x \to x(M)$. Aus (4) folgt $e \to e(M) \equiv 1$. Somit ist $1 \equiv e(M) \in A$. Es seien M_1 und M_2 zwei verschiedene maximale Ideale. Dann gibt es ein Element x_0, das zu M_1, aber nicht zu M_2 gehört. Folglich ist

$$x_0(M_1) = f_{M_1}(x_0) = 0, \quad x_0(M_2) = f_{M_2}(x_0) \neq 0.$$

Also trennt A die Punkte von $\mathscr{M}$. Ferner ist die symmetrische Algebra A nach Definition invariant gegenüber der Konjugiertenbildung. Die Anwendung des Satzes von Stone-Weierstrass führt zum Beweis von (c).

Damit ist der Satz bewiesen.

A.4.3. Der Satz von Wiener. Aufgaben. Es gibt zahlreiche Anwendungen der Theorie der Banachschen Algebren. Wir erinnern an einige Resultate der Algebra und der Analysis, die im Laufe der früheren Betrachtungen erhalten wurden.

Eine Banachsche Algebra über $\mathbf{C}$, die zugleich ein Körper ist, ist isometrisch-isomorph zu $\mathbf{C}$.

Das Spektrum eines beliebigen beschränkten Operators in einem Banachraum ist nicht leer.

Für einen beliebigen beschränkten Operator A in einem Banachraum X existiert der Limes $\lim\limits_{n\to\infty} \sqrt[n]{\|A^n\|} = r(A)$. Das Spektrum von A liegt innerhalb eines Kreises um den Nullpunkt mit dem Radius $r(A)$.

Unter Verwendung der Theorie der kommutativen Banachschen Algebren beweisen wir jetzt folgenden Satz von WIENER:

Ist $x(\theta)$ nirgends 0 und kann man $x(\theta) = \sum\limits_{k=-\infty}^{\infty} x_k e^{ik\theta}$ in Form einer absolut konvergenten Fourierreihe darstellen, so läßt sich auch $y(\theta) = 1/x(\theta)$ in Form einer absolut konvergenten Fourierreihe darstellen.

Beweis. Wir betrachten die Algebra l_1 der absolut konvergenten Folgen (Beispiel 4 aus A.1.2.). Wir bestimmen den zu l_1 gehörenden Raum $\mathcal{M}$. Es genügt, einen Homomorphismus von l_1 in C zuerst für die Funktion $x_0(t) = e^{it}$ zu betrachten und ihn anschließend in eindeutiger Weise auf l_1 auszudehnen. Wir setzen $f_M(x_0) = f_M(e^{it}) = \zeta$. Dann ist $f_M(x_0^{-1}) = f_M(e^{-it}) = \zeta^{-1}$. Aus (2) folgt $|\zeta| = |f_M(x_0)| \leqq \|x_0\| = 1$, $|1/\zeta| = |f_M(x_0^{-1})| \leqq \|x_0^{-1}\| = 1$. Also ist $|\zeta| = 1$ und $\zeta = e^{i\theta}$. Auf diesem Wege erhält man eine eineindeutige Zuordnung zwischen $\mathcal{M}$ und dem Punkt $|\zeta| = 1$ auf dem Rand des Einheitskreises. Da die Funktion $x(\theta) = \sum\limits_k x_k e^{ik\theta}$ für $-\pi \leqq \theta < \pi$ nicht verschwindet, gehört die Folge $x = (\ldots, x_{-n}, \ldots, x_0, \ldots, x_n, \ldots)$ zu keinem maximalen Ideal. Die Folgerung aus Lemma 1 in A.3.2. zeigt dann, daß diese Folge in der Algebra l_1 invertierbar ist. Wir setzen $y = x^{-1} = (\ldots, y_{-n}, \ldots, y_0, \ldots, y_n, \ldots)$. Folglich ist

$$y(M) = f_M(y) = \sum_{k=-\infty}^{\infty} y_k e^{ik\theta} = f(x^{-1}) = \frac{1}{f(x)} = \frac{1}{\sum\limits_{k=-\infty}^{\infty} x_k e^{ik\theta}}.$$

Damit ist der Satz bewiesen.

Zwei weitere wichtige Anwendungen der Theorie der Banachschen Algebren, den Spektralsatz für beschränkte Operatoren und den Satz von STONE-ČECH, formulieren wir später als Aufgaben.

Aufgaben

1. (a) Man zeige, daß der Raum der maximalen Ideale der Algebra $\mathcal{A}$ (Beispiel 3 aus A.1.2.) eineindeutig und stetig auf die Punkte des Einheitskreises $|z| \leqq 1$ abgebildet werden kann.

(b) Man zeige, daß $\mathcal{A}$ regulär und symmetrisch ist (und somit ein Radikal besitzt, das nur aus dem Nullelement besteht).

2. Woran liegt es, daß l_1 (Beispiel 4 aus A.1.2.) nicht isometrisch-isomorph zum Raum $C_{\mathcal{M}}$, d. h. zum Raum der stetigen Funktionen auf $|\zeta| = 1$ ist?

3. Man beweise folgenden Satz: Es sei

$$x(z) = \sum_{k=0}^{\infty} x_k z^k, \quad \sum_{k=0}^{\infty} |x_k| < \infty, \quad x(z) \neq 0 \text{ für } |z| \leqq 1.$$

Dann kann man $y(z) = \frac{1}{x(z)}$ in Form einer Taylorreihe darstellen, die für $|z| \leqq 1$ absolut konvergiert.

4. Mit $C^n[a, b]$ bezeichnen wir die im Intervall $[a, b]$ n-mal stetig differenzierbaren Funktionen $x(t)$.

(a) Man zeige, daß $C^n[a, b]$ eine Banachsche Algebra bezüglich der üblichen Operationen und der Norm

$$\|x\| = \sum_{k=0}^{n} \frac{1}{k!} \max_{a \leqq t \leqq b} |x^{(k)}(t)|$$

ist.

(b) Man bestimme die maximalen Ideale von $C^n[a, b]$ (vgl. [17], S. 19, 20).

(c) Man zeige, daß $C^n[a, b]$ eine symmetrische Algebra mit dem Radikal $R = \{0\}$ ist. Wie steht es in diesem Fall mit der Anwendung von Satz 2?

5. Die Algebra der stetigen komplexwertigen Funktionen von beschränkter Variation auf dem Intervall $[0, 1]$ wird mit $CBV\,[0, 1]$ bezeichnet. Die Norm ist

$$\|x\| = \sup_{0 \leqq t \leqq 1} |x(t)| + \operatorname{Var}(x, [0, 1]).$$

(a) Man zeige, daß $CBV[0, 1]$ eine Banachsche Algebra ist.

(b) Man bestimme die maximalen Ideale dieser Algebra.

6. Man gebe eine Banachsche Algebra an, die mit ihrem Radikal übereinstimmt.

7. Man beschreibe alle abgeschlossenen Ideale der Algebra $C[a, b]$.

8. Es sei T ein vollständiger regulärer Raum im Sinne von 2.5.6. und B_T die Menge aller auf T definierten beschränkten komplexwertigen Funktionen mit den üblichen Operationen und mit der Norm $\|x\| = \sup_{t \in T} |x(t)|$.

(a) Man zeige, daß B_T eine reguläre symmetrische Algebra mit dem Radikal $R = \{0\}$ ist.

(b) Man zeige, daß man die Punkte von T homöomorph in den Raum $\mathscr{M}$ der maximalen Ideale der Algebra B_T einbetten kann, wobei das Bild von T bei dieser Einbettung eine in $\mathscr{M}$ dichte Untermenge ist.

(c) Man zeige, daß eine beliebige beschränkte komplexwertige Funktion auf dem Bild von T bei dieser Einbettung eine eindeutig bestimmte stetige Fortsetzung auf $\mathscr{M}$ besitzt.

Die Aussage (b) zusammen mit der Tatsache, daß $\mathscr{M}$ kompakt ist, ist der Inhalt des bekannten Satzes von Tichonov über Kompaktifizierung. Die Aussage (c) geht auf Stone und Čech zurück. Eine Kompaktifizierung, die die Eigenschaft (c) besitzt, nennt man maximal. (c) bedeutet, daß $\mathscr{M}$ eine maximale Kompaktifizierung ist (vgl. [13], IX.2).

9. Es sei H ein Hilbertraum. In der Algebra $\mathscr{L}(H, H)$ betrachten wir die kommutative Unteralgebra $B(A_0)$, die durch den selbstadjungierten Operator A_0 erzeugt wird (also durch den Abschluß der linearen Hülle der Potenzen von A_0).

(a) Man zeige, daß $B(A_0)$ regulär ist und für das Radikal $R = \{0\}$ gilt.

(b) Man zeige, daß $B(A_0)$ symmetrisch ist und

$$\overline{x(M)} = x^*(M)$$

gilt, wobei x^* der zum Operator $x \in B(A_0)$ adjungierte Operator und $x(M)$ die Abbildung aus A.4. ist. (Zur Herleitung von (b) vergleiche man auch Aufgabe 10 (c).)

Die Anwendung von Satz 2 aus A.4.2. auf die Algebra $B(A_0)$ führt zum sogenannten Spektralsatz für selbstadjungierte Operatoren (vgl. [13], IX, oder [70], II).

10. Eine (nicht notwendig kommutative) Banachsche Algebra heißt *Algebra mit Involution*, falls es eine Abbildung $X \to X$ mit den Eigenschaften

$$(x + y)^* = x^* + y^*, \quad (xy)^* = y^* x^*, \quad (\alpha x)^* = \bar{\alpha} x^*, \quad (x^*)^* = x$$

gibt. Algebren mit Involution nennt man *B^*-Algebren*, falls außerdem $\|xx^*\| = \|x\|^2$ gilt.

(a) Man zeige, daß $\mathscr{L}(H, H)$ eine B^*-Algebra ist (vgl. [13], IX.3).

(b) Man zeige, daß eine kommutative B^*-Algebra regulär ist (vgl. [13], IX.3).

(c) Man zeige, daß eine B^*-Algebra symmetrisch ist. Es gilt sogar $\overline{x(M)} = x^*(M)$ (vgl. [13], IX.3, Lemma von Arens).

Zusammen mit Satz 2 führen die Aussagen (b) und (c) zu Resultaten, die auf Gel'fand und Najmark zurückgehen und mitunter als Hauptsatz der Theorie der kommutativen Banachschen Algebren bezeichnet werden:

Eine kommutative B^*-Algebra ist isometrisch-isomorph zur Algebra $C_{\mathscr{M}}$. Hierbei gilt $\overline{x(M)} = x^*(M)$.

Es zeigt sich somit, daß man ein abstraktes algebraisches Objekt, das durch 24 Axiome beschrieben wird (13 Axiome für eine kommutative Algebra, 5 Axiome, die mit der Norm verknüpft sind, ein Vollständigkeitsaxiom und 5 Axiome für B^*-Algebren), in Gestalt einer Algebra aller stetigen Funktionen über einem kompakten Hausdorffraum realisieren kann.

Dieses Resultat erlaubt von einem einheitlichen Standpunkt aus, weit voneinander entfernt liegende Fakten zu betrachten, wie den Satz von Wiener über absolut konvergente trigonometrische Reihen, den Satz über die Spektralzerlegung selbstadjungierter Operatoren, die Kompaktifizierungssätze von Tichonov, Stone und Čech sowie weitere Aussagen.

Literaturverzeichnis[1])

[1] N. I. Achieser, I. M. Glasmann, Theorie der linearen Operatoren im Hilbert-Raum, 5. Aufl., Akademie-Verlag, Berlin 1968 (Übersetzung aus dem Russischen).
[2] P. S. Alexandroff, Einführung in die Mengenlehre und die Theorie der reellen Funktionen, 4. Aufl., VEB Deutscher Verlag der Wissenschaften, Berlin 1967 (Übersetzung aus dem Russischen).
[3] В. И. Авербух, О. Г. Смолянов, Теория дифференцирования в линейных топологических пространствах. Успехи матем. наук 22: 6 (138) (1967), 200—260.
[4] S. Banach, Théorie des opérations linéaires, Warszawa 1932.
[5] Ю. М. Березанский, Разложение по собственным функциям самосопряженных операторов, „Наукова думка", Киев 1965.
[6] S. Bochner, Vorlesungen über Fouriersche Integrale, Akademische Verlagsgesellschaft, Leipzig 1932. [Russische Übersetzung: С. Бохнер, Лекции об интегралах Фурье, Физматгиз, Москва 1962.]
[7] N. Bourbaki, Eléments de mathématique, Livre I: Théorie des ensembles, Hermann, Paris 1954—1956 [Russische Übersetzung: Н. Бурбаки, Теория множеств, „Мир", Москва 1965.]
[8] N. Bourbaki, Eléments de mathématique, Livre III: Topologie générale, Hermann, Paris 1949—1958. [Russische Übersetzung: Н. Бурбаки, Общая топология, Основные структуры, Физматгиз, Москва 1958.]
[9] N. Bourbaki, Eléments de mathématique, Livre V: Espaces vectoriels topologiques, Hermann, Paris 1953—1955. [Russische Übersetzung: Н. Бурбаки, Топологические векторные пространства, И. Л., Москва 1959.]
[10] M. M. Day, Normed Linear Spaces, 3. Aufl., Springer-Verlag, Berlin—Heidelberg—New York 1973 [Russische Übersetzung: М. М. Дэй, Линейные нормированные пространства, И. Л., Москва 1961.]
[11] J. Dieudonné, Grundzüge der modernen Analysis, 2. Aufl., Verlag Vieweg & Sohn, Braunschweig/ VEB Deutscher Verlag der Wissenschaften, Berlin 1972 (Übersetzung aus dem Englischen).
[12] N. Dunford, J. T. Schwartz, Linear Operators I, Interscience Publishers, New York—London 1958. [Russische Übersetzung: Н. Данфорд, Дж. Шварц, Линейные операторы (Общая теория), И. Л., Москва 1962.]
[13] N. Dunford, J. T. Schwartz, Linear Operators II, Interscience Publishers, New York—London 1963. [Russische Übersetzung: Н. Данфорд, Дж. Шварц, Линейные операторы (Спектральная теория), „Мир", Москва 1966.]
[14] R. E. Edwards, Functional Analysis, Holt, Rinehart and Winston, New York—Chicago—San Francisco—Toronto—London 1965. [Russische Übersetzung: Р. Эдвардс, Функциональный анализ, „Мир", Москва 1969.]

[1]) Die mit * versehenen Titel wurden bei der deutschen Ausgabe hinzugefügt.

[15] A. FRAENKEL, Abstract Set Theory, 2nd ed., North-Holland Publ. Compl., Amsterdam 1961.

[16] A. A. FRAENKEL, Y. BAR-HILLEL, Foundations of Set Theory, North-Holland Publ. Comp., Amsterdam 1958. [Russische Übersetzung: А. Френкель, И. Бар-Хиллел, Основания теории множеств, „Мир“, Москва 1966.]

[17] I. M. GELFAND, D. A. RAIKOW, G. E. SCHILOW, Kommutative normierte Algebren, VEB Deutscher Verlag der Wissenschaften, Berlin 1964 (Übersetzung aus dem Russischen).

[18] I. M. GELFAND, G. E. SCHILOW, Verallgemeinerte Funktionen (Distributionen) I, 2. Aufl., VEB Deutscher Verlag der Wissenschaften, Berlin 1967 (Übersetzung aus dem Russischen).

[19] I. M. GELFAND, G. E. SCHILOW, Verallgemeinerte Funktionen (Distributionen) II, 2. Aufl., VEB Deutscher Verlag der Wissenschaften, Berlin 1969 (Übersetzung aus dem Russischen).

[20] I. M. GELFAND, G. E. SCHILOW, Verallgemeinerte Funktionen (Distributionen) III, VEB Deutscher Verlag der Wissenschaften, Berlin 1964 (Übersetzung aus dem Russischen).

[21] I. M. GELFAND, N. J. WILENKIN, Verallgemeinerte Funktionen (Distributionen) IV, VEB Deutscher Verlag der Wissenschaften, Berlin 1964 (Übersetzung aus dem Russischen).

[22] И. Ц. Гохберг, М. Г. Крейн, Введение в теорию линейных несамосопряженных операторов, „Наука“, Москва 1965.

[23] И. Ц. Гохберг, М. Г. Крейн, Теория вольтерровых операторов в гильбертовом пространстве и ее приложения, „Наука“, Москва 1967.

[24] P. R. HALMOS, Measure Theory, Van Nostrand, New York 1950. [Russische Übersetzung: П. Халмош, Теория меры, И. Л., Москва 1953.]

[25] P. R. HALMOS, Finite Dimensional Vector Spaces, 2. ed., Van Nostrand, Princeton (N. J.) 1958. [Russische Übersetzung: П. Халмош, Конечномерные векторные пространства, Физматгиз, Москва 1963.]

[26] P. R. HALMOS, Lectures on Ergodic Theory, The Math. Society of Japan 1956. [Russische Übersetzung: П. Халмош, Лекции по эргодической теории, И. Л., Москва 1959.]

[27]* P. R. HALMOS, A Hilbert Space Problem Book. Van Nostrand Comp., Princeton (N.J.) 1957. [Russische Übersetzung: П. Халмош, Гильбертово пространство в задачах, „Мир“, Москва 1970.]

[28] E. HILLE, R. S. PHILLIPS, Functional analysis and semi-groups (Revised edition), Amer. Math. Soc. Colloq. Publ. 31, Providence (R. I.) 1957. [Russische Übersetzung: Е. Хилле, Р. С. Филлипс, Функциональный анализ и полугруппы, И. Л., Москва 1962.]

[29] Л. В. Канторович, Функционалый анализ и прикладная математика, Успехи матем. наук 3: 6 (28) (1948), 89—185.

[30] L. W. KANTOROWITSCH, G. P. AKILOW, Funktionalanalysis in normierten Räumen, Akademie-Verlag, Berlin 1964 (Übersetzung aus dem Russischen).

[31]* T. KATO, Perturbation Theory for Linear Operators, Springer-Verlag, Berlin—Heidelberg—New York 1966. [Russische Übersetzung: Т. Като, Теория возмущений линейных операторов, „Мир“, Москва 1972.]

[32] J. KELLEY, General Topology, Van Nostrand, Princeton (N.J.) 1955. [Russische Übersetzung: Дж. Келли, Общая топология, „Наука“, Москва 1968.]

[33]* D. KLAUA, Allgemeine Mengenlehre, 1. Ein Fundament der Mathematik, 2. Aufl., Akademie-Verlag, Berlin 1968.

[34]* G. KÖTHE, Topologische lineare Räume I, Springer-Verlag, Berlin—Göttingen—Heidelberg 1960.

[35] М. А. Красносельский, Топологические методы в теории нелинейных интегральных уравнений, Гостехиздат, Москва 1956.

[36]* A. KUFNER, J. KADLEC, Fourier Series, Academia, Prague 1971.

[37] C. KURATOWSKI, Topologie I, 4. Aufl., Warszawa 1958. [Russische Übersetzung: К. Куратовский, Топология I, „Мир“, Москва 1966.]

[38] А. Лебег (H. LEBESGUE), Интегрирование и отыскание примитивных функций, ГТТИ, Москва 1934.

[39]* L. A. LJUSTERNIK, W. I. SOBOLEW, Elemente der Funktionalanalysis, 4. Aufl., Akademie-Verlag, Berlin 1968 (Übersetzung aus dem Russischen).

[40] M. LOÈVE, Probability Theory, Van Nostrand Comp., Toronto—NewYork—London 1955. [Russische Übersetzung: М. Лоэв, Теория вероятностей, И. Л., Москва 1962.]

[41] L. LOOMIS, An Introduction to Abstract Harmonic Analysis, Van Nostrand, New York 1955. [Russische Übersetzung: Л. Люмис, Введение в абстрактный гармонический анализ, И. Л., Москва 1956.]

[42] K. MAURIN, Methods of Hilbert Spaces, PWN, Warszawa 1967.

[43] S. G. MICHLIN, Vorlesungen über lineare Integralgleichungen, VEB Deutscher Verlag der Wissenschaften, Berlin 1962 (Übersetzung aus dem Russischen).

[44]* S. G. MICHLIN, Lehrgang der mathematischen Physik, Akademie-Verlag, Berlin 1972 (Übersetzung aus dem Russischen).

[45]* P. H. MÜLLER, Höhere Analysis I, II, Akademie-Verlag, Berlin 1972.

[46] I. P. NATANSON, Theorie der Funktionen einer reellen Veränderlichen, 3. Aufl., Akademie-Verlag, Berlin 1969 (Übersetzung aus dem Russischen).

[47] M. A. NEUMARK, Normierte Algebren, VEB Deutscher Verlag der Wissenschaften, Berlin 1959 (Übersetzung aus dem Russischen; 2. russische Auflage: М. А. Наймарк, Нормированные кольца, „Наука", Москва 1968).

[48] M. A. NEUMARK, Lineare Differentialoperatoren, Akademie-Verlag, Berlin 1960 (Übersetzung aus dem Russischen; 2. russische Auflage: М. А. Наймарк, Линейные дифференциальные операторы, „Наука", Москва 1969).

[49] R. PALEY, N. WIENER, Fourier Transforms in the Complex Domain. Amer. Math. Soc. Colloq. Publ., New York 1934. [Russische Übersetzung: Н. Винер, Р. Пэли, Преобразование Фурье в комплексной области, „Наука", Москва 1964.]

[50] А. И. Плеснер, Спектральная теория линейных операторов, „Наука", Москва 1965.

[51] F. RIESZ, B. SZ.-NAGY, Vorlesungen über Funktionalanalysis, 3. Aufl., VEB Deutscher Verlag der Wissenschaften, Berlin 1973 (Übersetzung aus dem Französischen).

[52] A. P. ROBERTSON, W. J. ROBERTSON, Topological Vector Spaces, Cambridge Univ. Press, Cambridge 1964. [Russische Übersetzung: А. П. Робертсон, В. Дж. Робертсон, Топологические векторные пространства, „Мир", Москва 1967.]

[53] W. RUDIN, Principles of mathematical analysis, 2. ed., McGraw-Hill Book Comp., New York—San Francisco—Toronto—London 1964. [Russische Übersetzung: У. Рудин, Основы математического анализа, „Мир", Москва 1966.]

[54] S. SAKS, Theory of the Integral, Warszawa—Lwów 1937. [Russische Übersetzung: С. Сакс, Теория интеграла, И. Л., Москва 1949.]

[55]* H. H. SCHAEFER, Topological Vector Spaces, MacMillan Comp., New York 1966. [Russische Übersetzung: X. Шефер, Топологические векторные пространства, „Мир", Москва 1971.]

[56]* L. SCHWARTZ, Méthodes mathématique pour les sciences physique. Hermann, Paris 1961. [Russische Übersetzung: Л. Шварц, Математические методы для физических наук, „Мир", Москва 1965.]

[57]* L. SCHWARTZ, Analyse mathématique I, II, Hermann, Paris 1967. [Russische Übersetzung: Л. Шварц, Анализ, „Мир", Москва 1972.]

[58] L. SCHWARTZ, Théorie des distributions, Hermann, Paris 1973 (Erstauflage 1950/51).

[59]* S. L. SOBOLEW, Einige Anwendungen der Funktionalanalysis auf Gleichungen der mathematischen Physik, Akademie-Verlag, Berlin 1964 (Übersetzung aus dem Russischen).

[60] Г. Е. Шилов, Математический анализ, Второй специальный курс, Физматгиз, Москва 1965.

[61] Г. Е. Шилов, Б. Л. Гуревич, Интеграл, мера и производная, Общая теория, „Наука", Москва 1967.

[62] Г. Е. Шилов, Фан Дык Тинь, Интеграл, мера и производная на линейных пространствах, „Наука", Москва 1967.

[63] E. C. Titchmarsh, Introduction to the Theory of Fourier Integrals, 2. ed., Oxford Univ. Press, Oxford 1948. [Russische Übersetzung: Е. Тичмарш, Введение в теорию интегралов Фурье, Гостехиздат, Москва 1948.]

[64] F. G. Tricomi, Integral Equations, Interscience Publ. New York—London 1957. [Russische Übersetzung: Ф. Трикоми, Интегральные уравнения, И. Л., Москва 1960.]

[65]* H. Triebel, Höhere Analysis, VEB Deutscher Verlag der Wissenschaften, Berlin 1972.

[66] Н. Я. Виленкин и другие. Функциональный анализ (серия „Справочная матем. библиотека"), „Наука", Москва 1964. [Nachauflage: М. Ш. Бирман и другие, „Наука", Москва 1972.]

[67] Б. З. Вулих, Теория полуупорядоченных пространств, Физматгиз, Москва 1961.

[68] N. Wiener, The Fourier Integral and Certain of Its Applications, Cambridge 1933. [Russische Übersetzung: Н. Винер, Интеграл Фурье и некоторые его приложения, Физматгиз, Москва 1963.]

[69] W. S. Wladimirow, Gleichungen der mathematischen Physik, VEB Deutscher Verlag der Wissenschaften, Berlin 1972. [Übersetzung aus dem Russischen. Zweite erweiterte russische Auflage: В. С. Владимиров, Уравнения математической физики, „Наука", Москва 1971.]

[70] K. Yosida, Functional Analysis, 3. Aufl., Springer-Verlag, Berlin—Göttingen—Heidelberg 1971. [Russische Übersetzung: К. Иосида, Функциональный анализ, „Мир", Москва 1967.]

[71] A. Zygmund, Trigonometric Series, 2nd ed., Cambridge Univ. Press, Cambridge 1959. [Russische Übersetzung: А. Зигмунд, Тригонометрические ряды I, II, „Мир", Москва 1965.]

Verteilung der Literatur auf die einzelnen Kapitel

1: 2, 7, 16, 33, 45, 57, 58.
2: 2, 4, 8, 11, 32, 34, 37, 55, 57.
3: 4, 9 bis 12, 14, 17, 19, 21, 25, 27 bis 31, 34, 39, 41, 42, 47, 51, 52, 65 bis 67, 70.
4: 1, 4, 5, 10, 12, 13, 15, 17 bis 23, 31, 41, 42, 44, 47, 48, 50, 56, 59, 60, 65, 69, 70.
5: 12, 24, 26, 38, 40, 45, 46, 49, 51, 54, 57, 61, 62.
6: 12, 30.
7: 38, 46, 51, 54, 61.
8: 6, 18 bis 21, 36, 49, 60, 68, 71.
9: 43, 64.
10: 3, 11, 28, 30, 35, 53

Anhang: 12, 13, 17, 70.

Sachverzeichnis

